ANALYSIS AND MODELING OF PARASITIC EFFECTS IN ADVANCED SILICON-ON-INSULATOR CMOS TECHNOLOGIES, INCLUDING NONCLASSICAL ULTRA-THIN-BODY TRANSISTORS

By

JI-WOON YANG

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2004

ACKNOWLEDGMENTS

I would like to express my sincere appreciation to the chairman of my supervisory committee, Professor Jerry G. Fossum, for his guidance and support throughout the course of this work. His great knowledge in semiconductor physics motivated my devotion to the field of semiconductor devices. He was a role model for me, put things in proper perspective, and contributed to my positive attitude. I would also like to thank the members of my supervisory committee (Professors Gijs Bosman, William R. Eisenstadt, and Timothy J. Anderson) for their guidance and interest in this work. I appreciate Erlinda Lane for all of her help preparing trips for numerous research reviews and conferences.

I am grateful to the Semiconductor Research Corporation and the University of Florida for their financial support. I thank Freescale Semiconductor Inc. for providing much of the data and information in this work.

I would also like to thank fellow students Meng-Hsueh Chiang, Keunwoo Kim, Mario M. Pelella, Lixin Ge, Kehuey Wu, Vishal Trivedi, Murshed Chowdhury,Weimin Zhang, Seung-Hwan Kim, Dong-Jun Yang, Seong-Mo Yim, Sang-Hoon Choi, Jang-sup Yoon, Hyeopgoo Yeo, Byung-Seo Kim, and Byung-Il Koh for their insightful and technical discussions

and friendships. Also, I thank all of the friends who made my years at the University of Florida such an enjoyable chapter of my life.

I must express heartfelt thanks to my father, Jeong-Yoon Yang, and my mother, Kil-Ja Kim, for their endless love and support in many ways through the years. Finally, I thank my wife, Kyung-Sun Lee, and my lovely daughter, Munhee Yang, for thier constant encouragement, support, help, sacrifice, prayer, and love throughout my studies.

In closing, I thank God. For only through God's grace and blessings has this pursuit been possible.

TABLE OF CONTENTS

page

ACKNOWLEDGMENTS ii

KEY TO ABBREVIATIONS vi

ABSTRACT viii

CHAPTER

1 INTRODUCTION 1

2 GATE-TO-BODY DIRECT TUNNELING CURRENT: MODELING AND IMPLEMENTATION 9

- 2.1 Introduction 9
- 2.2 Direct Gate Tunneling Current Components 10
- 2.3 Inversion-Region Model 15
  - 2.3.1 Valence-Band Electron Tunneling 15
  - 2.3.2 Scaling Effects 22
  - 2.3.3 Impact-Ionization Effects 28
- 2.4 Depletion/Accumulation-Region Model 34
- 2.5 Implementation of Gate-to-Body Direct Tunneling Current in UFPDB 39
- 2.6 Conclusions 49

3 ASSESSMENT OF THE GATE-TO-BODY DIRECT TUNNELING-CURRENT EFFECTS IN FLOATING-BODY PARTILLY DEPLETED SOI CMOS DEVICES AND CIRCUITS 50

- 3.1 Introduction 50
- 3.2 Pass-Gate Transistor and the Transient BJT Current 51
- 3.3 Pass Transistor-Based Wide Multiplexer 53
- 3.4 Dynamic Logic 55
  - 3.4.1 Dynamic 4-Way OR Circuit 55
  - 3.4.2 Pseudo Two-Phase Dynamic 2-Way NAND Circuit 57
  - 3.4.3 Dynamic Cascade Voltage Switch Logic Circuit 59
- 3.5 Sense Amplifier 64

3.6 Conclusions 70

4 PARASITIC BIPOLAR JUNCTION TRANSISTOR IN NONCLASSICAL ULTRA-THIN-BODY TRANSISTORS: MODELING, VERIFICATION, AND APPLICATION 71

4.1 Introduction 71
4.2 Effects of Device Structure on Gate-Induced Drain Leakage Current 72
4.3 Modeling of the Parasitic BJT 76
4.4 Model Implementation/Verification 89
4.5 Model Application 97
4.6 Conclusions 101

5 TRIPLE-GATE MOSFET: INSIGHT, FEASIBILITY, AND ANALYSIS 102

5.1 Introduction 102
5.2 Suppression of Corner Effects in TGFETs 106
5.3 Short-Channel Effects of Scaled TGFETs 113
5.4 Layout Area 120
5.5 Conclusions 126

6 SUMMARY AND SUGGESTIONS FOR FUTURE WORK 128

6.1 Summary 128
6.2 Suggestions for Future Work 131

APPENDIX DERIVATION OF THE INTERBAND TUNNELING CURRENT EQUATION 133

REFERENCE LIST 136

BIOGRAPHICAL SKETCH 144

KEY TO ABBREVIATIONS

BJT	bipolar junction transistor
CMOS	complementary metal-oxide-semiconductor
CVSL	cascade voltage switch logic
DG	double-gate
DGFET	double-gate MOSFET
DIBL	drain-induced barrier lowering
DOS	density-of-states
FB	floating body
FD	fully depleted
FDFET	planar fully-depleted MOSFET
GIDL	gate-induced-drain leakage
HP	high performance
IC	integrated circuit
LSTP	low standby power
MOSFET	metal-oxide-semiconductor field-effect transistor
PD	partially depleted
SCE	short-channel effect
SOI	silicon-on-insulator
STI	shallow trench isolation
TG	triple-gate

TGFET	triple-gate MOSFET
UFDG	University of Florida double-gate (model)
UFPDB	University of Florida partially depleted SOI and bulk MOSFET (models)
UTB	ultra-thin body
VLSI	very-large-scale integration

Abstract of Dissertation Presented to the Graduate
School of the University of Florida in Partial
Fulfillment of the Requirements for the Degree of
Doctor of Philosophy

ANALYSIS AND MODELING OF PARASITIC EFFECTS IN ADVANCED SILICON-ON-INSULATOR CMOS TECHNOLOGIES, INCLUDING NONCLASSICAL ULTRA-THIN-BODY TRANSISTORS

By

Ji-Woon Yang

August 2004

Chair: Jerry G. Fossum
Major Department: Electrical and Computer Engineering

This dissertation mainly focuses on analysis and modeling of parasitic effects in advanced silicon-on-insulator (SOI) complementary metal-oxide-semiconductor (CMOS) technology. As the channel length is reduced to the nanometer regime, the suppression of parasitic effects becomes an increasingly difficult technological challenge. First, the direct gate tunneling current is a major parasitic leakage component because extremely thin gate oxide is required in order to provide sufficient current drive with reduced supply voltage, and to control short-channel effects (SCEs). Gate-to-body tunneling current, which modulates the body voltage in scaled partially depleted (PD) SOI MOS field-effect transistors (FETs), is physically modeled and implemented in the University of Florida PD SOI/Bulk (UFPDB) MOSFET model. Several scaling effects (i.e., quantization effect in the channel, exchange energy of inversion electrons, and Fermi-Dirac statistics in the high-doped gate) are identified and accounted for. Using the model in UFPDB, it is shown how the tunneling current tends to suppress dynamic floating-body effects in some PD SOI circuits, providing benefits that follow from the complex interactions among the bias conditions, circuit topologies, and switching patterns. The parasitic bipolar-junction-transistor (BJT) effects in the floating-body SOI MOSFET are modeled because they can induce transient BJT current and amplify the gate-induced-drain-leakage (GIDL) current in nonclassical devices. The previous version of the BJT model in the University of Florida double-gate (UFDG) MOSFET model is modified and upgraded to describe correctly the majority-carrier concentration, which governs the parasitic BJT gain, in the (undoped) ultra-thin body (UTB) of nonclassical MOSFETs. The triple-gate MOSFET (TGFET), proposed as a means to alleviate the stringent UTB processing requirements, can be undermined by the top corners of the silicon-fin body, which can significantly affect the channel current-voltage characteristics of the TGFET. It is shown that the corner effects can be eliminated by leaving the body undoped. Then, for viable undoped bodies, the UTB dimensions needed to control SCEs in nanoscale TGFETs are examined by three-dimensional numerical device simulations. The results show that much more stringent body scaling is needed for undoped TGFETs relative to the doped ones, which are technologically and electrically infeasible. When the undoped-body dimensions are scaled for adequate SCE control, it is found that the TGFET suffers from a significant layout-area disadvantage relative to the double-gate (DG) FinFET and the planar fully-depleted MOSFET (FDFET), and thus it is not feasible. It is concluded that the DG FinFET, which is more scalable than the FDFET, is the nonclassical device with the most potential for future nanoscale CMOS applications.

CHAPTER 1
INTRODUCTION

Scale-down of complementary metal-oxide-semiconductor (CMOS) device dimensions in conventional bulk-silicon CMOS technology has been a primary driver over the past three decades in order to attain continued improvement in integrated-circuit (IC) performance and reduction in size (for lower cost). Silicon-on-insulator (SOI) technology has emerged as a competitor with bulk-silicon technology for high performance and as an effective means of extending CMOS scaling beyond bulk-silicon technology limits. Now, the partially depleted (PD) SOI MOS-field-effect-transistor (MOSFET), which is already in product stage, is rapidly replacing the bulk-silicon MOSFET for high-performance applications; and, with ultra-thin-bodies (UTB), the fully depleted SOI MOSFET and multi-gate MOSFETs are most promising candidates for scaling beyond the limit of conventional bulk-silicon CMOS technology.

As the channel length is reduced to the nanometer regime, however, the suppression of parasitic effects becomes an increasingly difficult technological challenge. Figure 1.1 shows the several parasitic effects of classical and nonclassical CMOS device structures. First, the direct gate-tunneling current can be a major parasitic leakage component

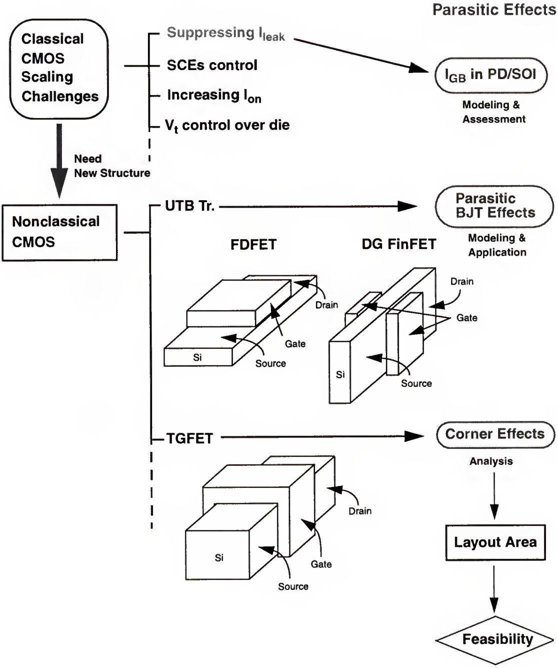


Figure 1.1 The several parasitic effects of classical and nonclassical CMOS device structures, with the implied works done in this dissertation indicated. Also, the structures of the primary nonclassical CMOS devices are illustrated.

because extremely thin gate oxide is required in order to provide sufficient current drive with reduced supply voltage, and to control short-channel effects (SCEs). Of the several components of the direct gate-tunneling current, the gate-to-body tunneling current modulates the body voltage and can influence the floating-body effects in scaled PD SOI MOSFETs. Therefore, the modeling of the gate-to-body direct tunneling current and assessment of its effects are crucial for the design of floating-body (FB) PD SOI circuits. The parasitic bipolar-junction-transistor (BJT) effects of the FB SOI MOSFET are also important because they induce transient BJT current and amplify the gate-induced-drain-leakage (GIDL) current. Thus, modeling of the parasitic BJT in undoped UTB devices illustrated in Figure 1.1, the most viable ones because of technological doping issues, is required to simulate precisely the dynamic and static off-state leakage currents in view of transient circuit operation and power consumption. The triple-gate MOSFET (TGFET) illustrated in Figure 1.1, which is proposed as a means to alleviate the stringent thin silicon-fin requirements of UTB devices, is subject to excessive conduction in the top corners of the silicon-fin, which can significantly affect the channel current-voltage characteristics of the TGFET. Analysis of the parasitic corner effects is needed to examine the feasibility of the TGFET.

As CMOS technology is scaled to the nanometer gate-length regime for higher density and speed, thinner gate oxide is required in order to provide sufficient current drive with reduced supply voltage, and to control short-channel effects. Sub-100nm CMOS technologies will require the (equivalent) gate oxide thickness to approach ~1nm [Sem01]. As the actual thickness (t_{ox}) is decreased, the direct gate-tunneling current increases with nearly exponential dependence on it. The higher gate current not only increases the standby power consumption of a highly integrated CMOS circuit, but also can adversely affect MOSFET performance. Therefore, previous studies of direct tunneling currents in MOS devices have focused primarily on the predominant component which flows between the gate and the channel [Reg99, Yan99, Kha00, Lo97, Yan00]; the much smaller tunneling current that flows between the gate and the body has not been investigated extensively [Maj00, Sha99, Lee00], nor modeled properly. However, in PD SOI MOSFETs, the latter current modulates the body voltage and can influence the floating-body effects [Fun00, Jos01, Chu02, Poi02]. Such influence, which can be beneficial or detrimental, can prevail even for relatively thick values of t_{ox} for which the larger gate-channel current is negligible. In chapter 2, we analyze the gate-to-body direct tunneling current based initially on Harrison's independent-electron tunneling model [Har61]. However, for the inversion case, the proper characterization of the gate-to-body tunneling current is somewhat different from Harrison's model since the electron density-of-states effective mass is different in the valence and conduction bands, which contain the initial and final states of the tunneling electrons that constitute the main component of the current. Moreover, several scaling effects (i.e., quantization in the channel, exchange energy of inversion electrons, and Fermi-Dirac carrier distribution in the poly-Si gate) and impact ionization, which have been overlooked in contemporary modeling of gate tunneling current, are identified and accounted for. We generalize the presentation by discussing the several gate-to-body tunneling current components in both the nMOSFET and the pMOSFET, for both inversion and accumulation cases, after which we describe our physics-based analytic modeling of the predominant components. The model is implemented in our process/ physics-based compact model (UFPDB [Fos02a]), which is unified for PD SOI and bulk-silicon MOSFETs.

The FB effects in PD SOI MOSFETs, and the resulting hysteresis and instability in dynamic operations, pose major challenges for very-large-scale-integration (VLSI) circuit designs. Further, the gate-to-body tunneling current modulates the body voltage, which is varied by capacitive coupling in transient operation, and thus can influence the FB effects [Fun00, Jos01, Chu02, Poi02]. Such influence can be beneficial or detrimental for the dynamic operations of FB PD SOI circuits and must be rigorously examined by a reliable physical model. In chapter 3, using the gate-to-body current model in UFPDB/Spice3, we assess the significance of the gate-to-body current for circuit topologies and switching patterns susceptible to FB effects.

The BJT effects of the FB SOI MOSFET have constituted a critical issue for a long time because they induce a premature breakdown [You88, Cho91], latch [Che88, Cho91], transient BJT current [Pel95], and also amplify the GIDL current [Che92]. While impact ionization-induced phenomena (premature breakdown and latch) have become mitigated due to scaled operation voltage, the transient BJT current and the amplified GIDL current are still major concerns for nonclassical UTB devices regarding control of off-state leakage current in view of dynamic and static power consumption [Sak03, Fos99, Sch04]. However, the parasitic BJT model [Kri96] of UFDG [Fos03a], which is our process/physics-based compact model for UTB devices, is inadequate to describe the transient BJT action and the amplification of GIDL current in nonclassical UTB devices because it does not correctly account for accumulation in the undoped UTBs. In chapter 4, the previous BJT model is modified and upgraded to correctly describe the majority-carrier concentration in the UTB, which governs the parasitic BJT gain. Further, circuit application with the upgraded UFDG in Spice3 [Fos03a] is exemplified by simulations of a nonclassical dynamic logic circuit, which show the parasitic BJT effect on circuit operation and stress the need to account for it in IC design at future CMOS technology nodes.

Conventional scaling of classical bulk-silicon and PD SOI CMOS has become very challenging because control of SCEs for acceptable I_{on}/I_{off} ratios requires precise channel doping levels and gradients [Fra01] that are, ultimately, impossible to achieve. As noted, promising alternatives for continued CMOS scaling are the nonclassical devices having UTBs, i.e., the planar fully depleted single-gate SOI MOSFET (FDFET) [Dor02] and the quasi-planar double-gate (DG) FinFET [Ked02], which give good

SCE control via the UTB. However, the good SCE control in the nonclassical nanoscale devices requires that the UTB thickness (t_{Si}) be considerably thinner than the gate length (L_g), as well as being uniform across the device and circuit. The TGFET was recently proposed [Doy03a, Doy03b] as a means to alleviate the stringent thin-t_{Si} requirements of UTB FDFETs and DG FinFETs. However, design studies of TGFETs [Doy03b] have revealed that the top corners of the silicon body can, with strong dependence on their shape, or radius of curvature, significantly affect the channel current-voltage characteristics of the TGFETs. Typically, the corner components of current reflect a lower threshold voltage (V_t) and higher off-state current (I_{off}) relative to the respective properties of the bulk TG device, and for some body structures they can cause a “hump” in the subthreshold log(I_{DS})-V_{GS} characteristic [Doy03b]. Hence, a serious technological issue is implied. And, the layout area of the TGFET is an issue for nanoscale CMOS [Doy03b]. The feasibility of the TGFET is thus not yet clear. In chapter 5, we examine the abnormal corner effects in nanoscale TGFETs, using two-dimensional (2-D) numerical device simulations and quasi-2-D analysis to gain physical insights on how the effects could possibly be suppressed, irrespective of the radius of curvature, which is difficult to control technologically. Then, we examine the feasibility of the TGFET for nanoscale CMOS applications with regard to SCEs and layout area. We check the SCEs for varying dimensions of the body, with and without doping, using a three-dimensional (3-D) numerical device simulator. Then, with the SCEs adequately controlled via body scaling in accord with the 3-D simulation results, we analyze the gate layout area of the integrated TGFET needed for current drive, and compare it with those of the FDFET and DG FinFET.

CHAPTER 2
GATE-TO-BODY DIRECT TUNNELING CURRENT: MODELING AND IMPLEMENTATION

2.1 Introduction

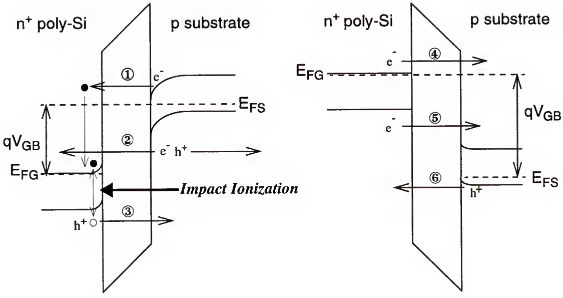
As CMOS technology is scaled to the nanometer gate-length regime for higher density and speed, thinner gate oxide is required in order to provide sufficient current drive with reduced supply voltage, and to control SCEs. Sub-100nm CMOS technologies will require the (equivalent) gate oxide thickness to approach ~1nm [Sem01]. As the actual thickness (t_{ox}) is decreased, the direct gate-tunneling current increases with nearly exponential dependence on it. The higher gate current not only increases the standby power consumption of a highly integrated CMOS circuit, but also can adversely affect MOSFET performance. Therefore, previous studies of direct tunneling currents in MOS devices have focused primarily on the predominant component which flows between the gate and the channel [Reg99, Yan99, Kha00, Lo97, Yan00]; the much smaller tunneling current that flows between the gate and the body has not been investigated extensively [Maj00, Sha99, Lee00], nor modeled properly. However, in PD SOI MOSFETs, the latter current modulates the body voltage and can influence the floating-body effects [Fun00, Jos01, Chu02, Poi02]. Such influence, which can be beneficial or detrimental, can prevail even for relatively thick values of t_{ox} for which the larger gate-channel current is negligible.

In this chapter, we analyze the (poly-Si) gate-to-body direct tunneling current based initially on Harrison's independent-electron tunneling model [Har61]. However, for the inversion case, the proper characterization of the gate-to-body tunneling current is somewhat different from Harrison's model since the electron density-of-states effective mass is different in the valence and conduction bands, which contain the initial and final states of the tunneling electrons that constitute the main component of the current. Moreover, several scaling effects (i.e., quantization in the channel, exchange energy of inversion electrons, and Fermi-Dirac carrier distribution in the poly-Si gate) which have been overlooked in contemporary modeling of gate tunneling current are identified and accounted for. We generalize the presentation by discussing the several gate-to-body tunneling current components in both the nMOSFET and the pMOSFET, for both inversion and accumulation cases, after which we describe our physics-based analytic modeling of the predominant components. The model is implemented in our process-based compact model (UFPDB [Fos02a]), which is unified for PD SOI and bulk-silicon MOSFETs.

2.2 Direct Gate Tunneling Current Components

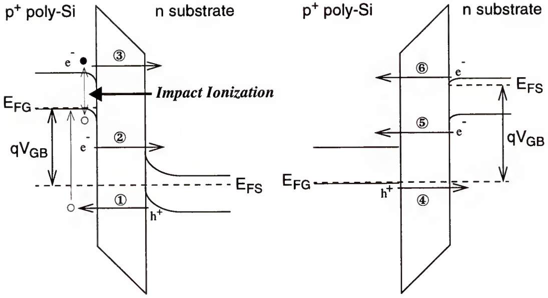
The components of gate tunneling current are identified using the simple energy-band diagrams of the n^+poly-Si/SiO_2/p-Si nMOSFET and p^+poly-Si/SiO$_2$/n-Si pMOSFET structures in Figure 2.1 and Figure 2.2, respectively. For the nMOSFET biased to inversion ($V_{GB} > V_t$) in Figure 2.1(a), the conduction-band electron tunneling (①) is the largest component because the associated Si-SiO$_2$ barrier height (3.1eV) is small compared to those (4.2eV and 4.5eV, respectively) associated with the valence-band electron tunneling (②) and the valence-band hole tunneling (③). However, even though the valence-band tunneling components are relatively small, they can significantly influence FB effects in PD SOI nMOSFETs because they result in hole injection in the body. The latter component ③ is seemingly very small because of the negligible concentration of holes in the valence band of the gate. However, as we suggest herein, it becomes predominant at high V_{GB} because of the impact ionization (illustrated in the figure) driven by hot electrons in the gate from component ①, which enhances the hole density. Therefore, all three tunneling components in Figure 2.1(a) can be important in this case.

In transient operation of the SOI nMOSFET, the dynamic V_{GB} can be negative due to the capacitive coupling between the body and the source/drain. Therefore, the accumulation case in Figure 2.1(b) must be considered. In this case, all three gate tunneling current components contribute to floating-body effects since the source and drain are isolated. The conduction-band electron tunneling (④) is the dominant component since the associated Si-SiO$_2$ barrier height is relatively small. The impact ionization driven by hot electrons in the substrate and hot holes in the



(a) Inversion (b) Accumulation

Figure 2.1 Energy bands and components of direct gate tunneling current in the nMOSFET with n^+ poly-Si gate for the inversion and accumulation cases. The impact-ionization effect is indicated for the inversion case. E_{FS} and E_{FG} represent the Fermi-level energy (E_F) of the silicon and the gate, respectively.



(a) Inversion (b) Accumulation

Figure 2.2 Energy bands and components of direct gate tunneling current in the pMOSFET with p^+ poly-Si gate for the inversion and accumulation cases. The impact-ionization effect is indicated for the inversion case.

gate, from components ④ and ⑥, respectively, is not significant because the hole and electron densities, respectively, in the substrate valence band and the gate conduction band are high for this bias condition. Therefore, the tunneling component ④ is the only important one in this case.

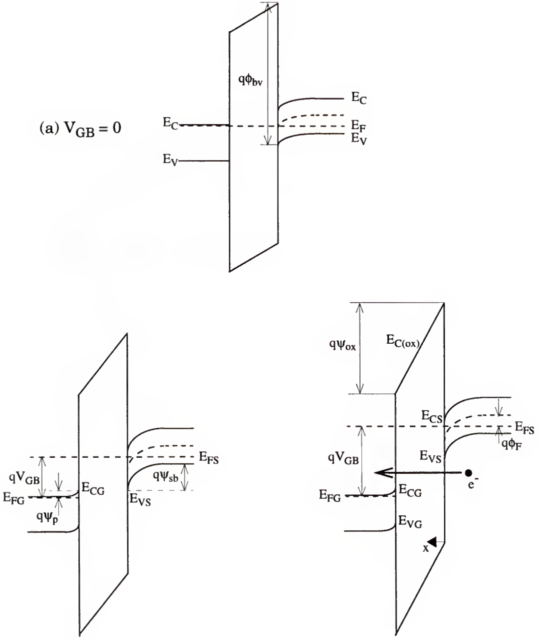
For the pMOSFET biased to inversion, all tunneling components in Figure 2.2(a) are analogous to those of the nMOSFET in Figure 2.1(a), and can be important. However, there is a difference in the accumulation case in Figure 2.2(b). Because of the small Si-SiO_2 barrier height of the conduction-band electron tunneling (⑥) relative to the valence-band hole tunneling (④), the component ⑥ becomes important for $V_{GB} > V_{FB}$ when the electron density in the substrate is substantial. Therefore, the gate-to-body direct tunneling current is mainly the component ⑥, whereas it is the component ④ for the nMOSFET in Figure 2.1(b).

In summary, the gate-to-body direct tunneling currents are mainly due to I_{VBE} (valence-band electron tunneling ②) and I_{VBH} (valence-band hole tunneling ③) at inversion, and I_{CBE} (conduction-band electron tunneling ④) at accumulation in the nMOSFET, and due to I_{VBE} (valence-band electron tunneling ②) and I_{CBE} (conduction-band electron tunneling ③) at inversion, and I_{CBE} (conduction-band electron tunneling ⑥) at accumulation in the pMOSFET. These tunneling currents modulate the threshold voltage by changing the body voltage in PD SOI MOSFETs, and therefore influence the FB effects [Fun00]. Herein we describe physics-based analytic modeling of the gate-to-body direct tunneling currents, and its implementation in our compact process-based model (UFPDB [Pel02]) for PD SOI and bulk-silicon MOSFETs.

2.3 Inversion-Region Model

2.3.1 Valence-Band Electron Tunneling

We focus first on the direct valence-band electron tunneling, which is a major component of the gate-to-body current at the normal inversion mode of both the nMOSFET and the pMOSFET (before the onset of impact ionization due to the tunneling from the channel). As seen in Figure 2.1(a) and Figure 2.2(a), the mechanisms are exactly the same in both devices. Therefore, the model we develop here for the nMOSFET is directly applicable to the pMOSFET as well. The valence-band electron tunneling in the nMOSFET for $V_{GB} > 0$, i.e., component ② in Figure 2.1, as governed by the electric potential band bending ψ_{sb} in the silicon substrate and the potential drop ψ_p in the gate due to poly-Si depletion [Tau98]. Valence-band electrons in the substrate tunnel into the conduction band of the gate when E_{CG} (E_C at the poly-Si surface) in the gate drops below E_{VS} (E_V at the silicon surface) in the substrate, as depicted in Figure 2.3. Note that this bias condition implies strong inversion. The tunneling current density J (in x) is characterized by modifying the independent-electron tunneling model of Harrison [Har61] to account for the different transverse density-of-states (DOS) effective masses in the silicon valence band (m_{ta}^*) and in the poly-Si conduction band (m_{tb}^*). The more general expression for the net (silicon valence band



(b) $V_{GB} = V_{onset}$ $(E_{VS} = E_{CG})$

(c) $V_{GB} > V_{onset}$ $(E_{VS} - E_{CG} > 0)$

Figure 2.3 Energy bands showing the onset of the valence-band electron tunneling in the nMOSFET (with $V_{BS} = 0$).

(a) to poly-Si conduction band (b)) electron tunneling current density (magnitude) is given by (see Appendix)

$$J_{a\rightarrow b} = \frac{4\pi q}{h^3}\left[m_{ta}{}^*\int_0^{E_{t(max)}}\int_0^{\infty} Pf_a(1-f_b)dE_x dE_t \right.$$
$$\left. -m_{tb}{}^*\int_0^{E_{t(max)}}\int_0^{\infty} Pf_b(1-f_a)dE_x dE_t\right];(2.1)$$

$m_{ta}{}^*$, $m_{tb}{}^*$: transverse DOS effective masses in the states a and b, respectively;

E_x: energy (kinetic plus potential) in the x-direction (normal to the Si-SiO$_2$ interface);

E_t: transverse (parallel to the Si-SiO$_2$ interface) energy (kinetic) component;

f_a, f_b: probabilities (Fermi-Dirac) of occupation of states a and b, respectively;

P: tunneling probability.

Note that (2.1) simplifies to the commonly used direct tunneling current expression [Maj00, Sha99, She72, Dep95] when $m_{ta}{}^*$ and $m_{tb}{}^*$ are equal, e.g., in the case of conduction-band electron tunneling.

The model in (2.1) assumes conservation of E_t as well the total energy ($E = E_t + E_x$) in the (elastic) tunneling process, which, based on the WKB approximation for the matrix element for the transition [Har61] and a one-band parabolic dispersion relation involving an effective mass ($m_{ox}{}^*$) in the SiO$_2$ [Dep95], yields

$$P = \exp\left[-2\int_0^{t_{ox}} \sqrt{\frac{2m_{ox}^*}{\hbar^2}(E_{C(ox)} - E_x)dx}\right] \qquad (2.2)$$

where, from Figure 2.3, $E_{C(ox)}(x) = E_{VS} + \phi_{bv} - q\mathcal{E}_{ox}x$ is the conduction band-edge energy in the oxide; $\phi_{bv} = 4.2\text{eV}$ and $\mathcal{E}_{ox} = \psi_{ox}/t_{ox}$ is the (constant) electric field in the oxide. However, (2.1) and (2.2) do not explicitly account for the conservation of transverse momentum (via phonons) in this case of an indirect valence band-conduction band transition [Mol64]. This accounting decreases P dramatically, but the functional dependence on $\mathcal{E}_{ox}$ is almost identical [Mol64] to (2.2), but with m_{ox}^* being a larger effective mass [Sze81]. Several papers, e.g., [Dep95], report different electron effective masses in the oxide, corresponding to conduction band-conduction band tunneling and depending on t_{ox} and model assumptions. We infer from measurements of the valence-band tunneling current (discussed later) $m_{ox}^* = 0.36m_0$, where m_0 is the free electron mass.

Referring to Figure 3(c), we see that $f_a \cong 1$, and (as we have verified numerically) that we can assume (with <1% error) $f_b \cong 0$. Thus, for the tunneling component of interest, (2.1) simplifies to

$$J_{VBE} = \frac{4\pi q m_v^*}{h^3}\int_0^{E_{t(max)}}\int_0^{\infty} P dE_x dE_t \qquad (2.3)$$

where m_v^* (= m_{ta}^*) is the transverse DOS effective mass for electrons in the silicon valence band. We assume $m_v^*=0.65m_0$, which follows from summing the heavy ($m_{vh}^*=0.49m_0$) and light ($m_{vl}^*=0.16m_0$) hole masses

[Sze81] in the valence subbands to properly account for the composite 2D DOS. Note in Figure 3(c) that the total energy (E) of a tunneling electron contributing to the valence-band component of interest here must be between E_{CG} and E_{VS}; the tunneling to states in the gate below E_{CG} is negligible. We can thus define the E_x limits of integration in (2.3) as $E_{VS} - E_t$ and $E_{CG} - E_t$; and $E_{t(max)}$ is $E_{VS} - E_x$. However, E_x and E_t are correlated in (2.3), and hence we need a simplifying assumption to do the integration. We assume that in the valence band, electron kinetic energy is negligible compared to the potential energy, or $E_t \cong 0$. Therefore, (2.3) is approximated as

$$J_{VBE} = \frac{4\pi q m_v^*}{h^3}\int_{E_{CG}}^{E_{VS}}\int_0^{E_{VS}-E_x} P\,dE_t\,dE_x = \frac{4\pi q m_v^*}{h^3}\int_{E_{CG}}^{E_{VS}}(E_{VS}-E_x)P\,dE_x . \quad (2.4)$$

To evaluate (2.4), (2.2) is first integrated for specific E_x:

$$P = \exp\left[-\frac{4t_{ox}\sqrt{2m_{ox}^*}}{3q\hbar}\,\frac{(q\phi_{bv}+E_{VS}-E_x)^{\frac{3}{2}}-(q\phi_{bv}+E_{VS}-E_x-q\psi_{ox})^{\frac{3}{2}}}{\psi_{ox}}\right]. \quad (2.5)$$

With (2.5) substituted in (2.4), we note that $(E_{VS} - E_x)$ is a much weaker function than the exponential function. Therefore, it can be removed from the integration, provided a correction factor is inserted. Numerical solution show that a factor of 1/2 is acceptable; thus,

$$J_{VBE} = \frac{4\pi m_v^* q}{h^3}\frac{(E_{VS}-E_{CG})}{2}$$

$$\int_{E_{CG}}^{E_{VS}} \exp\left[-\frac{4t_{ox}\sqrt{2m_{ox}^*}}{3q\hbar}\frac{(q\phi_{bv}+E_{VS}-E_x)^{\frac{3}{2}}-(q\phi_{bv}+E_{VS}-E_x-q\Psi_{ox})^{\frac{3}{2}}}{\Psi_{ox}}\right]dE_x, \tag{2.6}$$

which we rewrite as

$$J_{VBE} \cong \frac{A}{2}(E_{VS}-E_{CG})$$

$$\int_{E_{CG}}^{E_{VS}} \exp\left[-Bt_{ox}\frac{(q\phi_{bv}+E_{VS}-E_x)^{\frac{3}{2}}-(q\phi_{bv}+E_{VS}-E_x-q\Psi_{ox})^{\frac{3}{2}}}{\Psi_{ox}}\right]dE_x \tag{2.7}$$

$$A = \frac{4\pi m_v^* q}{h^3},\ B = \frac{4\sqrt{2m_{ox}^*}}{3q\hbar}.$$

Integrating (2.7), and noting that for inversion conditions

$$\frac{Bt_{ox}}{\Psi_{ox}}\left[q\phi_{bv}^{\frac{3}{2}}-(q\phi_{bv}-q\Psi_{ox})^{\frac{3}{2}}\right] \gg 1$$

$$\frac{Bt_{ox}}{\Psi_{ox}}\left[(q\phi_{bv}+E_{VS}-E_{CG})^{\frac{3}{2}}-(q\phi_{bv}+E_{VS}-E_{CG}-q\Psi_{ox})^{\frac{3}{2}}\right] \gg 1,$$

we obtain

$$J_{VBE} = \frac{A}{3B}\frac{t_{ox}}{\Psi_{ox}}E_d$$

$$\left\{F_1\exp\left[-Bt_{ox}\frac{q\phi_{bv}^{\frac{3}{2}}-(q\phi_{bv}-q\Psi_{ox})^{\frac{3}{2}}}{\Psi_{ox}}\right]-F_2\exp\left[-Bt_{ox}\frac{q\phi_{bv}'^{\frac{3}{2}}-(q\phi_{bv}'-q\Psi_{ox})^{\frac{3}{2}}}{\Psi_{ox}}\right]\right\} \tag{2.8}$$

$$F_1 = \left[q\phi_{bv}^{\frac{1}{2}} - (q\phi_{bv} - q\psi_{ox})^{\frac{1}{2}}\right]^{-1}, \; F_2 = \left[q\phi_{bv}'^{\frac{1}{2}} - (q\phi_{bv}' - q\psi_{ox})^{\frac{1}{2}}\right]^{-1}$$

$E_d = E_{VS} - E_{CG}$, $q\phi_{bv}' = q\phi_{bv} + E_d$.

From Figure 2.3, for $E_F = E_C$ in the highly doped poly-gate region, we can write

$$E_d = E_{VS} - E_{CG} = q\left[V_{GB} - \psi_{sb} - \psi_p - \frac{E_g}{2q} + \phi_F\right]; \tag{2.9}$$

E_g is the silicon bandgap and ϕ_F is the Fermi potential in the substrate. We note that J_{VBE} in (2.8) must be zero when E_d is zero or negative. To validate the tunneling current model then, E_d in (2.8) is replaced by the smoothing function,

$$\overline{E_d}/q = \frac{\ln\{1 + \exp[S_{VBE}(E_d/q)]\}}{S_{VBE}} \tag{2.10}$$

where S_{VBE} is a constant. For $E_d > 0$, $\overline{E_d} \cong E_d$; but when E_d decreases to zero and becomes negative, $\overline{E_d}$ approaches zero, which properly forces J_{VBE} to zero.

For a preliminary check of the J_{VBE} model defined by (2.8)-(2.10), we combine it with a numerical solution of the basic MOS equation [Tau98],

$$V_{GB} - \Phi_{PS} = \psi_{sb} + \psi_{ox} + \psi_p , \tag{2.11}$$

coupled to the 1-D (in x) Poisson equation; Φ_{PS} is the poly-Si/Si work-function difference, which is approximated by assuming that the Fermi level in the bulk gate region is pinned at $E_{C(gate)}$ due to the high doping density, and y_p is the potential drop in the poly-Si gate, which is characterized based on the depletion approximation [Tau98]. We show in Figure 2.4 model-predicted J_{VBE} for three different values of t_{ox} (1.65nm, 1.79nm, and 2.18nm, which were determined via ellipsometer measurements), with measured gate-to-body tunneling currents superimposed. The measured data were obtained from bulk-silicon nMOSFETs via a carrier-separation experiment as described in [Shi98]. The gate was positively biased (V_{GB}), while the source/drain and the substrate were grounded. The source/drain current ($J_{S/D}$) thereby measured is component ① in Fig. 1, while the measured substrate current (J_B) is component ②. The model predictions, with m_{ox}^* tuned to $0.38m_0$ and S_{VBE}=14.0, are in good agreement with the measured data for relatively low V_{GB}, but underestimate the gate-to-body tunneling current for high V_{GB}, especially for thicker gate oxide.

2.3.2 Scaling Effects

The measured data in Figure 2.4 were taken from scaled nMOSFETs. The discrepancies between the data and the J_{VBE} model predictions thus could possibly be due to scaled-device effects that we heretofore ignored. In this section, we consider three such effects: electron degeneracy in the gate and electron energy quantization and exchange energy in the inversion layer. The noted degeneracy, due to the high gate doping density, results in E_F being above E_C in the gate, with the difference $\eta = (E_F - E_C)/(k_BT)$ defined by Fermi-Dirac statistics [Sze81]:

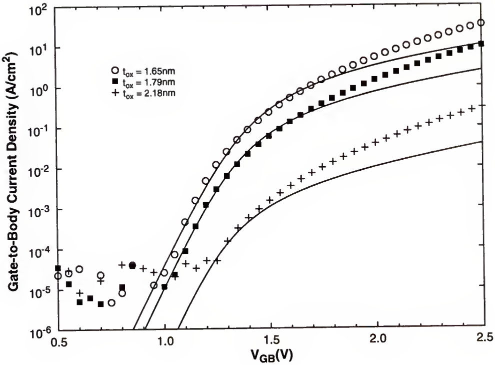


Figure 2.4 Measured (points) and modeled (curves) gate-to-body tunneling current-density (J_{VBE}) characteristics in nMOSFETs with varying gate-oxide thickness. The gate area is 10x10mm^2, the substrate doping density is 1.7×10^{17}cm^{-3}, and the n^+-poly-Si gate doping density is approximately 1×10^{20}cm^{-3}; $V_{BS} = V_{DS} = 0$, $m_{ox}^* = 0.38m_0$.

$$N_p = N_c F_{1/2}(\eta) \tag{2.12}$$

where N_p is the n^+poly-gate doping density and N_c is the effective density of states in the conduction band. (For the p^+ poly-Si gate on the pMOSFET, N_c is replaced by N_v, the effective density of states in the valence band.) We use Nilsson's approximation [Nil78] for η:

$$\eta \cong \ln\left(\frac{N_D}{N_c}\right) + \frac{N_D/N_c}{[64 + 3.6(N_D/N_c)]^{1/4}} \,. \tag{2.13}$$

In this approximation, $E_F - E_C$ is 57.0meV when N_D is 10^{20}cm^{-3}, which is significant. Thus, the substrate-gate overlapped energy for tunneling, (2.9), should be replaced by

$$E_d = E_{VS} - E_{CG} = q\left[V_{GB} - \psi_{sb} - \psi_p + V_T\eta - \frac{E_g}{2q} + \phi_F\right] \tag{2.14}$$

where $V_T = k_BT/q$ is the thermal voltage.

The noted quantization effect, for a specified V_{GB}, increases ψ_{sb} [Tau98], and hence can affect J_{VBE} [Yan99]. In the UFPDB model [Pel02], the quantization effect is implemented [Chi01] employing van Dort's model [Van94], where the effect is modeled by calculating an effective intrinsic carrier density corresponding to the bandgap widening $\Delta E_{g(QM)}$ implied by the creation of 2-D electron subbands above E_C. However, the noted exchange energy lowers the energy of the subbands, thereby effectively narrowing the bandgap. We use Stern's model [Ste73] to account for this narrowing:

$$\Delta E_{g(EX)} = \frac{q}{3\pi^2 \bar{k}\varepsilon_0}\left(\frac{2\pi N_i}{n_v}\right)^{\frac{1}{2}} \quad \text{(in eV)} \tag{2.15}$$

where N_i is the areal electron density ($N_i = -Q_i/q$ where Q_i is the inversion charge density approximated as $-\varepsilon_{Si}E_{xf0}$ [Tau98]; E_{xf0} is the surface electric field), n_v is the multiplicity of the valleys from which the lowest subband arises ($n_v = 2$ for inversion electrons, $n_v = 1$ for inversion holes), and $\bar{k}$ is the effective static dielectric constant, which is $(k_{Si} + k_{ox})/2$ in the extreme 2D electron-limit case. For strong inversion, we thus modify the bandgap widening to account for the exchange-energy narrowing:

$$\Delta E_g = \Delta E_{g(QM)} - \Delta E_{g(EX)} \; . \tag{2.16}$$

Now, for specified V_{GS} (= V_{GB} with V_{BS} = 0) with (2.16), the model-defined ψ_{sb} and Q_i (from UFPDB discussed in Sec. 2.5) are used with (2.8) and (2.14) to define J_{VBE}. For the t_{ox} = 1.65nm nMOSFET, we show in Figure 2.5 the model-predicted J_{VBE} with and without the three scaled-device effects, compared with the measured substrate current. Note that when the quantization effect is accounted for, both the degeneracy and exchange-energy effects must be modeled in order to get reasonable agreement with the data. This agreement is shown in Figure 2.6 for all values of t_{ox}. The better agreement between the measured and model-predicted J_{VBE} shows the importance of the noted scaled-device effects. For all three devices, a tuned value of $m_{ox}^* = 0.36m_0$, which is lower than the value inferred previously without the scaled-device effects, was used.

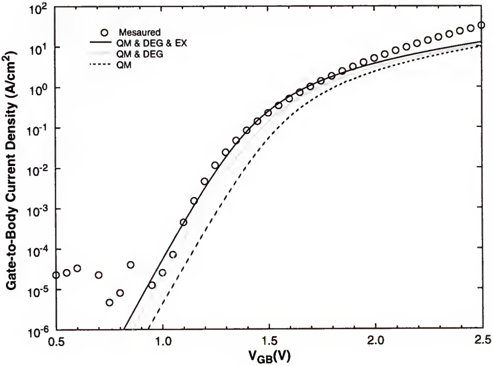


Figure 2.5 Modeled valence-band electron tunneling current density of the t_{ox}=1.65nm nMOSFET, showing the effects of quantization (QM), gate degeneracy (DEG), and channel exchange energy (EX); $V_{BS} = V_{DS} = 0$ and $m_{ox}^* = 0.36m_0$. The measured gate-to-body current is also shown.

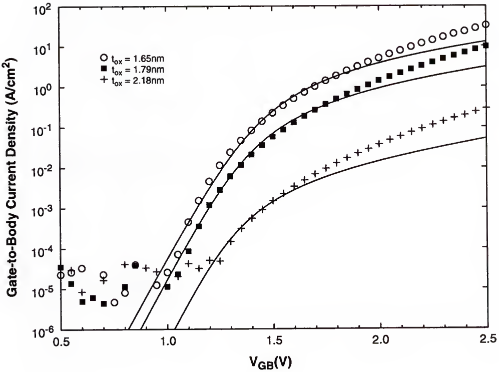


Figure 2.6 Measured and modeled gate-to-body tunneling current-density (J_{VBE} with scaled-device effects) characteristics of nMOSFETs with varying gate-oxide thickness; $V_{BS} = V_{DS} = 0$, $m_{ox}^* = 0.36m_0$.

The agreement is good for low V_{GB}, but large discrepancies remain for high V_{GB}.

2.3.3 Impact-Ionization Effects

The remaining model-measurement discrepancies evident in Figure 2.6 are similar to those reported in [Sha99] for thin t_{ox}, although the model-predicted curves in [Sha99] were tuned for better fits to the data at high V_{GB}. The discrepancies in [Sha99] were attributed to limitations of the WKB approximation for ultrathin oxides. However, we believe electron-hole pair generation by impact ionization in the gate and subsequent valence-band hole tunneling underlie the large discrepancies in Figure 2.6 and define the dominant gate-to-body direct current component at high V_{GB}.

As illustrated in Figure 2.1(a), the conduction-band electrons tunneling from the substrate, for high V_{GB}, become hot and have sufficient kinetic energy to drive significant impact ionization in the gate. A portion of the hot electrons will lose their energy via electron-electron interactions, but we presume some surface depletion in the poly-Si gate that reduces the electron density enough to enable a significant fraction of the electrons to lose their energy via impact ionization. Holes in the valence band generated by the impact ionization can then tunnel to the substrate valence band, thus significantly augmenting component ③ in Figure 2.1(a). To check this theory, we assume a steady-state condition where a hole density is established in the gate, and the hole tunneling current is defined directly by the rate of hole generation via the impact ionization. This assumption presumes that the tunneling process is much faster than the hole-electron recombination process in the gate, as well as that the hole tunneling without the impact ionization is negligible. Thus, with reference to Figure 2.3(c), we express the gate-substrate hole tunneling current as

$$J_{VBH} = \int_{E_{CS}}^{\infty} (M-1)dJ'_{CBE} \tag{2.17}$$

where J'_{CBE} is the mentioned fraction of the conduction band (channel)-gate electron tunneling current J_{CBE} (component ① in Figure 2.1(a)), E_{CS} is the conduction-band edge energy at the silicon surface, and (M-1) is the impact-ionization multiplication factor, which depends on the kinetic energy (E - E_{CG}) of the electrons in the gate; E is the total energy of the tunneling electrons. We express the kinetic energy in terms of the electron temperature [Sah91]:

$$E - E_{CG} = \frac{3}{2}k_B T_e . \tag{2.18}$$

For a first-order characterization of the impact ionization, we write

$$(M-1) \cong \alpha\lambda_e \tag{2.19}$$

where α, dependent on T_e, is the ionization rate and λ_e is the energy relaxation length. For quasi-steady-state conditions [Slo91],

$$\alpha(T_e) = \alpha_0 \exp\left[-\frac{\beta_0}{\gamma(T_e - T)}\right] \tag{2.20}$$

where $\gamma = 5k_B/(2q\lambda_e)$, α_0 and β_0 are constants, and T is the lattice temperature.

With (2.18)-(2.20), (2.17) with models for J_{CBE} and J'_{CBE} can be integrated over the conduction band to characterize J_{VBH}. To corroborate our theory, however, we assume $J'_{CBE} \sim J_{CBE}$ and use measurements of the conduction-band electron tunneling, which we assume occurs predominantly at or near $E_{CS} + \Delta E_g$ where ΔE_g is a effective bandgap widening in (2.16). Then, the integral in (2.17) is expressed as

$$\int_{E_{CS}}^{\infty} (M-1) dJ_{CBE} \cong (M-1) J_{CBE(meas)} \tag{2.21}$$

where (M-1) is evaluated at $E = E_{CS} + \Delta E_g$, yielding

$$J_{VBH} = \alpha\lambda_e J_{CBE(meas)} \tag{2.22}$$

with a evaluated from (2.20) at $T_e = 2(E_{CS} + \Delta E_g - E_{CG})/(3k_B)$, which follows from (2.18); the presumed kinetic energy can be expressed in terms of the gate voltage as evident in Figure 2.3(c):

$$E_{CS} + \Delta E_g - E_{CG} = E_d + E_g + \Delta E_g \tag{2.23}$$

where $E_d = E_{VS} - E_{CG}$ from (2.14). In (2.23), we have implicitly assumed that the transverse kinetic energy is negligible since it is not included in ΔE_g. (For high V_{DS}, the inversion electrons can have high transverse kinetic energy near the drain due to high electric field in channel. However, the gate-to-body direct tunneling current is negligible in this region, because of relatively low E_d and ε_{ox}, and hence the noted assumption is innocuous.) Figure 2.7 shows the kinetic energy in (2.23)

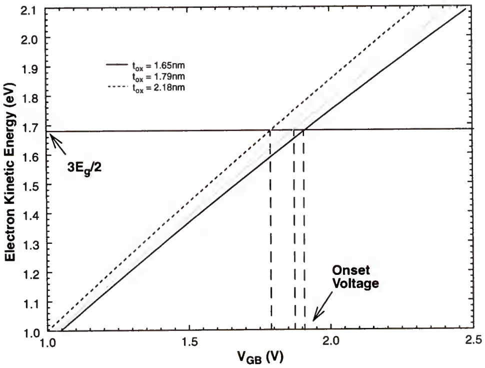


Figure 2.7 Simulated kinetic energy of the tunneling electron from the conduction band of the substrate. The onset voltage for impact-ionization is indicated for varying oxide thickness.

versus V_{GB}, derived via UFPDB, for the three values of t_{ox} in Figure 2.6. Our first-order characterization of J_{VBH} in (2.22) is applicable when the kinetic energy is greater than the physical onset value for impact ionization. We define this onset via a classical ballistic model governed by conservation of both energy and momentum [Sah91]; it estimates the onset energy as $3E_g/2$. Figure 2.7 shows that the onset V_{GB} increases with decreasing t_{ox} as implied by Figure 2.6.

For the t_{ox} = 1.65nm device of Figure 2.6, we plot in Figure 2.8 $J_{VBH}(V_{GB})$ evaluated from (2.22) for V_{GB} greater than the onset voltage (2.2V, as inferred from Figure 2.7). For the evaluations, we used the measured $J_{CBE(meas)}(V_{GB})$ ($= J_{S/D}$), also plotted in Figure 2.8, and we assumed $\alpha_0 = 2.45\times10^6\text{cm}^{-1}$ and λ_e = 65nm as given in [Slo91], and we tuned $\beta_0 = 2.98\times10^6\text{V-cm}^{-1}$, which is larger than, but comparable with the value given in [Slo91]. The larger β_0 could reflect some energy dissipation via electron-electron interaction in the gate. For comparison, we include in Figure 2.8 the measured gate-to-body direct tunneling current (J_B) and $J_{VBE}(V_{GB})$ from (2.8), which are in Figure 2.6, and the predicted sum ($J_{VB(total)}$) of J_{VBE} and J_{VBH}. Note how well $J_{VB(total)}(V_{GB})$ matches J_B for all V_{GB} in the inversion region, and that J_{VBH}, driven by the impact ionization, is the dominant component for high V_{GB}. Note also the steeper slope of $J_{VBH}(V_{GB})$, relative to that of $J_{VBE}(V_{GB})$ and $J_{CBE(meas)}(V_{GB})$. The V_{GB} dependence of (M-1) in (2.21) causes the steeper slope, which cannot be explained by possible uncertainties in t_{ox}, m_v^*, or m_{ox}^* in the

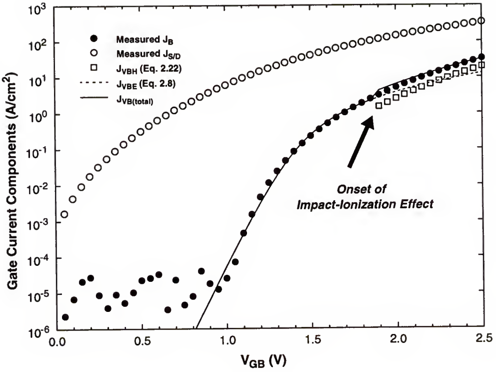


Figure 2.8 Measured and modeled gate-to-body current-density characteristics of an nMOSFET, including the two predicted valence-band components; $V_{BS} = 0$. The impact-ionization effect, driven by the conduction-band electron tunneling component, the measured characteristic of which is included, is illustrated.

$J_{VBE}(V_{GB})$ model while maintaining good agreement with the measured J_B for lower V_{GB}.

Using the same impact-ionization model with the same parameter values, we evaluated J_{VBH} and $J_{VB(total)}$ for the other devices in Figure 2.6 having different t_{ox}; the model predictions and the measured J_B are plotted versus V_{GB} in Figure 2.9. We used the onset voltages for impact ionization evident in Figure 2.7. The good agreement between the measured and model-predicted J_B at high V_{GB} in Figure 2.9 supports our theory of a significant impact-ionization effect. (Note that a smoothing function similar to that in (2.10) would have to be used in the model to effect the proper turn-on of the impact ionization.)

2.4 Depletion/Accumulation-Region Model

We next consider the gate-to-body tunneling for $V_{GB} < 0$ as reflected by the energy-band diagrams in Figure 2.10. For accumulation, the general tunneling current expression (2.1), with $m_{ta}^* = m_{tb}^* = m_c^*$ in the conduction band, describes the predominant conduction-band electron tunneling, i.e., component ④ in Figure 2.1(b), as

$$J_{CBE} = -\frac{4\pi q m_c^*}{h^3}\int_0^{E_{t(max)}}\int_0^{\infty} P(f_a - f_b)dE_x dE_t \qquad (2.24)$$

where $E_{t(max)}$ can be virtually infinite; m_c^* is the transverse DOS effective mass for electrons in the (100) silicon conduction band ($m_c^* = 0.19m_0$). Noting in Figure 2.10(b) that the minimum E_x must be E_{CG}, and

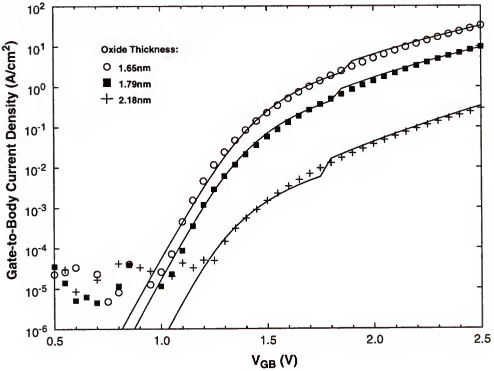
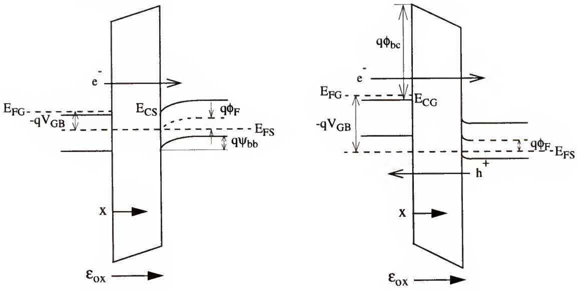


Figure 2.9 Measured and modeled gate-to-body current-density characteristics of nMOSFETs with varying gate-oxide thickness; $V_{BS} = V_{DS} = 0$, $m_{ox}=0.36m_0$, and $S_{VBE}=13.5$.



(a) Depletion

(b) Accumulation

Figure 2.10 Energy bands showing the conduction-band electron tunneling for the depletion and accumulation cases in the nMOSFET. The valence-band hole tunneling is also shown as a dominant component for strong accumulation in (b).

integrating (2.24), with the Fermi-Dirac distribution functions f(E = E_t + E_x), over E_t, we get

$$J_{CBE} = -\frac{4\pi q m_c^*}{h^3} kT \int_{E_{CG}}^{\infty} P(E_x) \ln\left(\frac{1 + \exp\left[-\frac{E_x - E_{FG}}{kT}\right]}{1 + \exp\left[-\frac{E_x - E_{FS}}{kT}\right]} \right) dE_x . \qquad (2.25)$$

A classical WKB approximation and a one-band parabolic dispersion relation was used to calculate the tunneling probability:

$$P = \exp\left[-2\int_0^{tox} k dx\right]. \qquad (2.26)$$

In Figure 2.10,

$$E_{COX}(x) = \phi_{bc} + E_{CG} + q\varepsilon_{OX}x \qquad (2.27)$$

where ε_{OX} is the electrical field strength in the oxide layer and ϕ_{bc} is the tunnel barrier height. Thus,

$$k = \frac{1}{\hbar}\sqrt{2m_{ox}^*(\phi_{bc} + E_{CG} - E_x + q\varepsilon_{OX}x)} \qquad (2.28)$$

and

$$P = \exp\left[-\frac{4t_{ox}\sqrt{2m_{ox}^*}}{3q\hbar} \frac{(\phi_{bc} + E_{CG} - E_x + q\psi_{OX})^{\frac{3}{2}} - (\phi_{bc} + E_{CG} - E_x)^{\frac{3}{2}}}{\psi_{OX}} \right] \qquad (2.29)$$

where $\psi_{ox} = \varepsilon_{ox} t_{ox}$.

In Figure 2.10(b),

$$E_x - E_{FG} = (E_x - E_{CG}) - k_B T\eta \qquad (2.30)$$

$$E_x - E_{FS} = (E_x - E_{FG}) - qV_{GB} \qquad (2.31)$$

where η (= $(E_{FG} - E_{CG}) / (k_BT)$) is the degeneracy due to the high gate doping density in (2.13). The integration in (2.26) is done numerically (trapezoidal) with a finite upper-E_x limit (E_{CG} + 0.3eV), partitioning the integration range into 50 equal segments.

As illustrated in Figure 2.10(a), a finite J_{CBE} actually flows prior to the onset of strong accumulation. The tunneling current flows to the channel, body, or both, depending on V_{GB}. Therefore, for modeling the gate-to-body component, we need the onset condition of gate-to-body current and it is defined by following smoothing function, which properly forces the gate-to-body current to zero when the surface potenital is above the some fraction of ϕ_F and a significant channel exists:

$$\overline{J_{CBE}} = J_{CBE} \times \left[1 + \exp\left(\frac{\psi_{sb} - F \times \phi_F}{S_{CBE}}\right)\right]^{-1} . \tag{2.32}$$

With (2.32), the formalism of the tunneling current in the accumulation bias region, (2.25), is applicable to calculate the tunneling current in the depletion bias region when the lower limit of the integral is E_{CS} instead of E_{CG}. However, we need to smooth the change of the lower-energy limit in each bias region. The following smoothing function is used:

$$\overline{E_{CG}} = E_{CS} + q \ln\left\{1 + \exp\left[S_{FB}\left(\frac{E_{CG} - E_{CS}}{q}\right)\right]\right\} / S_{FB} \tag{2.33}$$

where S_{FB} is a constant. For $E_{CG} < E_{CS}$, $E_{CG} \cong E_{CS}$; but when E_{CG} increases above E_{CS}, $E_{CG} \cong E_{CG}$, which properly change the lower limit.

For a preliminary check of our model for J_{CBE} in (2.25), we combine it with a numerical solution of the basic MOS equation (2.11), without ψ_p (i.e., without poly depletion for accumulation) but with the degeneracy effect in (2.13) included in Φ_{PS}, coupled to the 1-D (in x) Poisson equation. We show in Figure 2.11 the model-predicted J_{CBE} for three different values of t_{ox} (1.65nm, 1.79nm, and 2.18nm), with measured gate-to-body tunneling currents superimposed. The model predictions, with $m_{ox}^*=0.36m_0$, the same value finally used for the J_{VBE} model, are in good agreement with the measured data for relatively low $|V_{GB}|$ in the depletion/accumulation region, but underestimate the gate-to-body tunneling current (magnitude) for high $|V_{GB}|$. This difference is likely due to the valence-band hole tunneling, depicted in Figure 2.10(b), which becomes significant for strong accumulation of holes. Since normal operation of scaled CMOS devices would avoid this region, we do not include this component of tunneling current in our model. For the t_{ox} = 2.18nm device, the error seems to be prevalent even for lower $|V_{GB}|$, but we note in Figure 2.11 that this apparent discrepancy is probably due to the low values of J_{CBE} that are difficult to separate out from the noise in the measurement system.

2.5 Implementation of Gate-to-Body Direct Tunneling Current in UFPDB

The complete gate-to-body direct current model was implemented in UFPDB [Pel02, Chi01] for reliable prediction of floating-body effects in scaled PD/SOI CMOS devices and circuits. We overview in

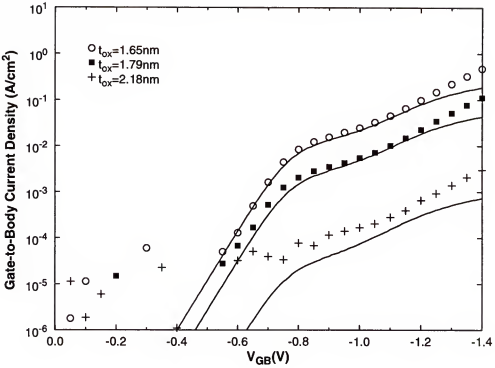


Figure 2.11 Measured (points) and modeled (curves) gate-to-body current-density (magnitude of $J_{CBE} < 0$) characteristics of nMOSFETs with varying gate-oxide thickness; $V_{DS} = 0$, $m_{ox}^* = 0.36m_0$, S_{CBE}=0.04, F=0.5, and S_{FB}=10.0.

this section the implementations of $I_{VBE}(V_{GS}, V_{DS})$ and $I_{CBE}(V_{GS}, V_{BS}, V_{DS})$.

In Sec. 2.3, $I_{VBE}(V_{GB}) \equiv I_{VBE}(V_{GS} = V_{GB}, V_{BS} = 0, V_{DS} = 0)$ was developed. For arbitrary bias, we write $V_{GB} = V_{GS} - V_{BS}$ and $\psi_{sb} = \psi_{sf} - V_{BS}$, where ψ_{sf}, used in UFPDB, is the absolute surface potential in the channel, referenced to that in the hypothetical unbiased body region. We note from (2.14) that we are hence using the $I_{VBE}(V_{GB})$ formalism directly with V_{GB} and ψ_{sb} replaced by V_{GS} and ψ_{sf}, respectively.

We extend the model to include V_{DS} dependence as follows. The total valence-band electron tunneling current is defined by integrating J_{VBE} along the channel:

$$I_{VBE} = W\int_0^{L_{eff}} J_{VBE}(\psi_{sf})dy \tag{2.34}$$

where L_{eff} is the effective channel length and W is the channel width. For strong inversion,

$$\psi_{sf}(y) = \psi_{sf0} + V(y) \tag{2.35}$$

where $\psi_{sf0} \equiv \psi_{sf}(y = 0)$ is the surface potential at the source end of the channel and V(y) is the quasi-Fermi potential for the inversion electrons. Assuming that the electron channel current is constant in y, we approximate V(y) as [Tau98]

$$V(y) = \frac{V_{GfS} - V_t}{m} - \left[\left(\frac{V_{GfS} - V_t}{m}\right)^2 - \frac{2y}{L_{eff}}\left(\frac{V_{GfS} - V_t}{m}\right)V_{DS} + \frac{y}{L_{eff}}V_{DS}^2\right]^{\frac{1}{2}} \tag{2.36}$$

where m = 1 + C_d/C_{ox}; C_{ox} is the gate-oxide capacitance, C_d is the body depletion capacitance, and V_t is a conventionally defined threshold voltage. Since UFPDB does not involve V_t directly, we assume in (2.36) that

$$\frac{V_{GfS} - V_t}{m} \cong V_{DS(sat)} \tag{2.37}$$

where $V_{DS(sat)}$ is the saturation drain voltage defined in UFPDB [Chi01]. Further, to smoothly account for saturation, we replace V_{DS} in (2.36) by V_{DSx}, a UFPDB-defined smoothing function that equals V_{DS} for low V_{DS}, but is limited to $V_{DS(sat)}$ at high V_{DS}.

To complete the characterization of in I_{VBE} in (2.34), we need to characterize $\psi_p(y)$ in (2.14). This is done in UFPDB [Chi01] via the depletion approximation in the poly-Si gate, which yields the depletion charge in terms of ψ_p:

$$Q_p = qN_Px_p = \sqrt{2\varepsilon_{Si}qN_P\psi_p} \ . \tag{2.38}$$

Applying Gauss' law to the gate-oxide interface gives

$$Q_p = \frac{\varepsilon_{ox}\psi_{ox}}{t_{ox}} = C_{ox}(V_{GS} - \Phi_{PS} - \psi_{sf} - \psi_p) \tag{2.39}$$

which, with (2.38), yields

$$\psi_p(y) = \frac{1}{C_{ox}^2}(qN_P\varepsilon_{Si} + C_{ox}^2(V_{GS} - \Phi_{PS} - \psi_{sf})$$

$$-\sqrt{qN_P\varepsilon_{Si}[qN_P\varepsilon_{Si} + 2C_{ox}^2(V_{GfS} - \Phi_{PS} - \psi_{sf})]}) \text{,(2.40)}$$

with $\psi_{sf}(y)$ given by (2.35).

Now, for a specified bias (V_{GS}, V_{DS}), the integrand J_{VBE} in (2.34) is evaluated via (2.8), (2.14), and (2.35)-(2.40), using the strong-inversion ψ_{sf0} determined in UFPDB [Chi01]. To properly incorporate the quantization effect and the exchange energy into the ψ_{sf0} calculation, a few Newton-Raphson iterations are required since those effects and ψ_{sf0} are correlated. For the first iteration, ψ_{sf0} is a strong-inversion guess based only on the device structure. The quantization effect, dependent on $\Delta E_{g(QM)}$ [Van94], and the exchange energy dependent on Q_i in (2.15) are calculated using ψ_{sf0}. For the subsequent iterations, an effective intrinsic carrier density corresponding to the bandgap widening ΔE_g in (2.16) yields the new ψ_{sf0} iterate. Generally, four iterations are used.

The integration in (2.34) is then done numerically (trapezoidal method), partitioning L_{eff} into 10 equal segments. Note that the temperature dependence of I_{VBE} is defined implicitly via $y_{sf}(T)$, and explicitly via $N_c(T)$, $E_g(T)$, $f_F(T)$, and V_T in (14). Figure 2.12 shows the simulated I_{VBE} characteristics of an nMOSFET with varying V_{GS} and V_{DS}. Note how the integrated tunneling current decreases substantively with increasing V_{DS} due mainly to (E_{VS} - E_{CG}) decreasing along the channel.

In the accumulation region, the integrated I_{CBE} can be expressed simply as

$$I_{CBE} = WL_{eff}J_{CBE}(\psi_{sb}) \tag{2.41}$$

since the surface potential is constant along the channel. In (2.41), ψ_{sb}, which is calculated by subtracting V_{BS} from the UFPDB-defined [Chi01]

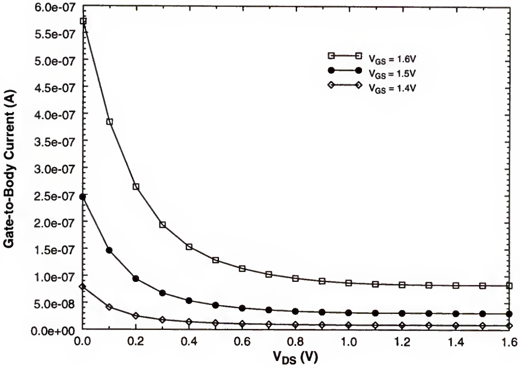


Figure 2.12 UFPDB-simulated valence-band electron tunneling current (I_{VBE}) characteristics of an nMOSFET for varying V_{GS} and V_{DS}; L_{eff} = W = 10mm, t_{ox} = 1.65nm, m_{ox}^* = 0.36m_0.

$\psi_{sf}(y = 0)$, is smoothed to zero when it tends to be less than zero in the accumulation region. The total gate-to-body current, I_{GB}, in UFPDB is defined by summing $I_{VBE} > 0$ in (2.34) and $I_{CBE} < 0$ in (2.41): $I_{GB} = I_{VBE} + I_{CBE}$.

Since the $I_{GB}(V_{GS}, V_{BS}, V_{DS})$ formalism is complex, we use difference approximations, based on ΔV_{GS} and ΔV_{DS} perturbations, in UFPDB to calculate the derivatives needed for the SPICE (Newton-Raphson) nodal analysis. The perturbed values of I_{VBE} follow directly from (2.34), with ψ_{sf0} virtually fixed because of the strong-inversion condition ($V_{GS} > V_{TS}$ [Chi01]). The electron effective mass in the oxide layer (m_{ox}^*) and the smoothing-function constant (S_{VBE}, S_{CBE}, and F) are the new model parameters needed. Further, m_{ox}^* is used as a flag to turn-on the I_{GB} model: if $m_{ox}^*=0$, then the I_{GB} formalism is bypassed (although the accounting for the exchange energy via (2.16) is permanent in UFPDB-2.0 for $V_{GfS} > V_{TS}$); S_{VBE} and S_{CBE} are used as flags to turn-on the I_{VBE} and I_{CBE} models, respectively.

For the pMOSFETs, with same values of t_{ox} as the nMOSFETs previously modeled, we show in Figure 2.13 the UFPDB-predicted J_{VBE}-V_{GB} characteristics. The model parameter m_{ox} is tuned as $0.37m_0$, versus $0.36m_0$ for the nMOSFETs. Since the tunneling processes in both devices are identical, the slightly different values of m_{ox} possibly reflects slight differences in the device structure, e.g., in t_{ox}. The model-measurement agreement in Figure 2.13 is very good for low and moderate V_{GB}, but large

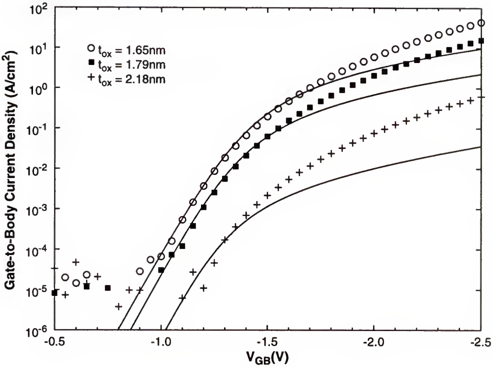


Figure 2.13 Measured and modeled (J_{VBE} with scaled-device effects) gate-to-body current characteristics of pMOSFETs; $V_{BS} = V_{DS} = 0$, $m_{ox} = 0.37m_0$.

discrepancies remain for high V_{GB} as we found for the nMOSFETs. Thus, the impact-ionization effect at high V_{GB} appears to occur in the pMOSFETs, as depicted in Figure 2.2(a), as well as in the nMOSFETs. As shown in Figure 2.13, the onset voltages of impact-ionization is slightly lower than those of nMOSFETs because the kinetic energy of tunneling electron in pMOSFETs expressed by (2.23) is bigger than that in nMOSFETs due to the higher degeneracy in p^+poly gate.

In Figure 2.14, we show the the UFPDB-predicted J_{VBH}-V_{GB} characteristics for the pMOSFETs. We don't include the conduction-band electron tunneling, which is dominant in deep accumulation region, since typical operation voltage in this scaled technology does not go to the region. The valence-band hole tunneling formalism is identical to the conduction-band electron tunneling in (2.25) except the transverse DOS effective mass for electrons in the silicon conduction-band (m_c*=0.19m_0), which should be replaced by the transverse DOS effective mass for electrons in the silicon valence band (m_v*=0.65m_0). The model predictions, with m_{ox}*=0.37m_0, which is the same value as in J_{VBE} modeling, S_{CBE}=0.045, F=0.8, and S_{FB}=15.0, are in good agreement with the measured data for depletion/accumualtion-region with relatively low V_{GB}.

Continuous scaling of the SiO_2 gate-dielectric thickness (t_{ox}) has resulted in significant direct gate tunneling current, as noted herein, not to mention the strong tendency for penetration of the poly-gate dopant into the silicon-channel region. Alternative gate dielectrics, having higher

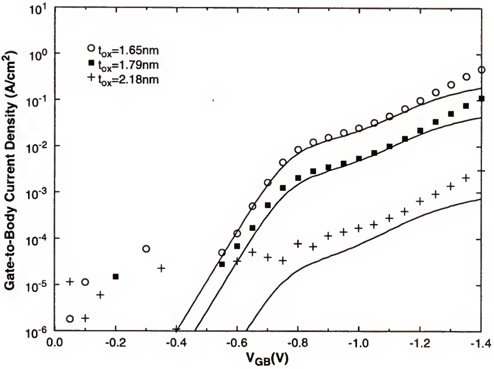


Figure 2.14 Measured and modeled (J_{VBH}) gate-to-body current-density characteristics of pMOSFETs with varying gate-oxide thickness; $V_{BS} = V_{DS} = 0$, $m_{ox} = 0.37m_0$, $S_{CBE} = 0.045$, $F = 0.5$, and $S_{FB} = 15.0$.

dielectric constants and hence enabling thicker t_{ox}, yet high gate capacitance and drive current, are thus being examined for ultra-scaled CMOS [Shi99]. Therefore, we further upgrade the UFPDB-2.0 model for arbitrary gate-dielectric permittivity, $\varepsilon_d = k_d\varepsilon_0$, where the k_d is the dielectric constant and ε_0 is the vacuum permittivity. In the upgraded code, $k_d\varepsilon_0$ replaces ε_{ox}, with k_d being a new model parameter.

2.6 Conclusions

Gate-to-body tunneling current in contemporary MOSFETs was physically modeled and implemented in the process-based compact model UFPDB [Pel02], which is unified for PD/SOI as well as bulk-silicon devices. The valence-band electron tunneling (J_{VBE}) to the n^+-poly-Si gate, predominant for inversion conditions, and the conduction-band electron tunneling (J_{CBE}) to the substrate, predominant for depletion/ accumulation conditions, were modeled by modifying the classical independent-electron formalism. For J_{VBE}, several scaling effects (i.e., quantization effect in the channel, exchange energy of inversion electrons, and Fermi-Dirac statistics in the high-doped gate), which have been overlooked in contemporary modeling of gate direct-tunneling current, were identified and accounted for. The composite I_{GB} model showed good agreement with measured data for gate-oxide thicknesses varying down to 1.65nm.

CHAPTER 3
ASSESSMENT OF THE GATE-TO-BODY DIRECT TUNNELING-CURRENT EFFECTS IN FLOATING-BODY PARTILLY DEPLETED SOI CMOS DEVICES AND CIRCUITS

3.1 Introduction

The FB effects in PD SOI MOSFETs, and the resulting hysteresis and instability in dynamic operations, pose major challenges for very-large-scale-integration (VLSI) circuit designs. Further, the gate-to-body tunneling current modulates the body voltage, which is varied by capacitive coupling in transient operation, and thus can influence the FB effects [Fun00, Jos01, Chu02, Poi02]. Such influence can be beneficial or detrimental for the dynamic operations of FB PD SOI circuits and must be rigorously examined by a reliable physical model. The dynamic V_{GB} can be positive or negative in dynamic operation of FB PD SOI devices and circuits, which is why I_{GB} had to be modeled (in chapter 2) in all inversion, depletion, and accumulation conditions.

In this chapter, we use an I_{GB} model in UFPDB/Spice3 to assess the significance of $I_{GB}(t)$ for circuit topologies and switching patterns susceptible to FB effects. For the device/circuit simulations, the UFPDB model was loosely calibrated to a contemporary scaled device with physical gate length equal to of 60nm, oxide thickness of 1.4nm, and retrograded channel with surface doping at $6.0\text{x}10^{17}\text{cm}^{-3}$. No special source/drain engineering for carrier recombination/generation control was assumed.

3.2 Pass-Gate Transistor and the Transient BJT Current

First, we consider a pass-gate transistor, in which the FB effect tends to drive the transient BJT current [Pel95], possibly resulting in excessive power consumption, degradation of noise margin and stability, and logic-state error in pass-gate and other dynamic-logic circuits [Chi96, Lu97, Chu98]. We assess the effect of reduced transient BJT current due to the gate-to-body current [Sin02]. Consider the FB PD SOI nMOSFET with source/drain initially biased at V_{DD} and the gate grounded in a steady-state condition as shown in Figure 3.1. This means that $V_{GS} = -V_{DD}$, with $V_{BS} = 0$, and hence holes are accumulated in floating body. When we allow for gate-to-body tunneling current, V_{BS} becomes negative due to body discharging via conduction-band electron tunneling (I_{CBE}); that is, the body voltage is now less than V_{DD}. Now, consider the source voltage switching to ground abruptly. The body voltage is capacitively coupled down by the ratio of gate capacitance and body-source junction capacitance, and then the transient $V_{BS}(t)$ becomes positive, thereby driving the transient BJT current. Since, with I_{GB}, the body voltage is initially less than V_{DD}, $V_{BS}(t)$ is decreased relative to that without I_{GB}, and the transient BJT current is reduced as shown by the UFPDB/Spice3 simulation result in Figure 3.1, by about 50% in this example with V_{DD} = 1.0V.

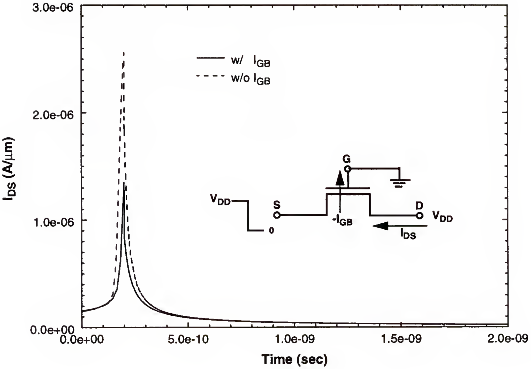


Figure 3.1 UFPDB/Spice3-predicted transient BJT current in the PD/SOI pass transistor (nMOSFET: L_{gate} = 60nm, t_{ox} = 1.4nm) shown, with and without the gate-to-body tunneling current ($I_{GB} < 0$) accounted for. The source-pulse fall time is 0.2ns, with V_{DD} = 1.0V.

Previous work concerning PD SOI CMOS pass-gate circuits [Lu97, Chu98] has shown that the transient BJT current can be problematic and necessitate modifications in circuit design to control its effects. However, as the CMOS technology is scaled, the reduced V_{DD} tends to alleviate the BJT effect as implied by the discussion concerning Figure 3.1. Hence, in scaled PD SOI CMOS pass-gate circuits, we do not expect the gate-to-body tunneling current to provide any significant benefit, even though it can reduce the transient BJT current as in Figure 3.1.

3.3 Pass Transistor-Based Wide Multiplexer

To check the I_{GB} effect on pass transistor-based logic, we examine the pass transistor-based n-to-1 multiplexer circuit in Figure 3.2(a), where n=32, operating at V_{DD} = 1.0V. We consider the worst-case scenario for the transient BJT effect as in [Lu97]. All inputs are assumed at "1" to start with. The control signal for N32 and P32 pair is "1", and hence the $V_{in,32}$ passes "1" to node-1, which continues to be "1" afterwards. All the control signals (from N1 and P1 pair to N31 and P31 pair) are "0", and the ($V_{in,1}$ - $V_{in,31}$) are then pulled down. As a result, transient BJT currents flow through the 31 unselected pairs to pull down node-1, which is supposed to be at "1" and is held by the single selected pair (N32 and P32). Unlike the previous report [Lu97], which shows the transient BJT current pulls down node-1 to the threshold voltage of the output buffer and can lead to a wrong logic state, the UFPDB/Spice3 simulation results in

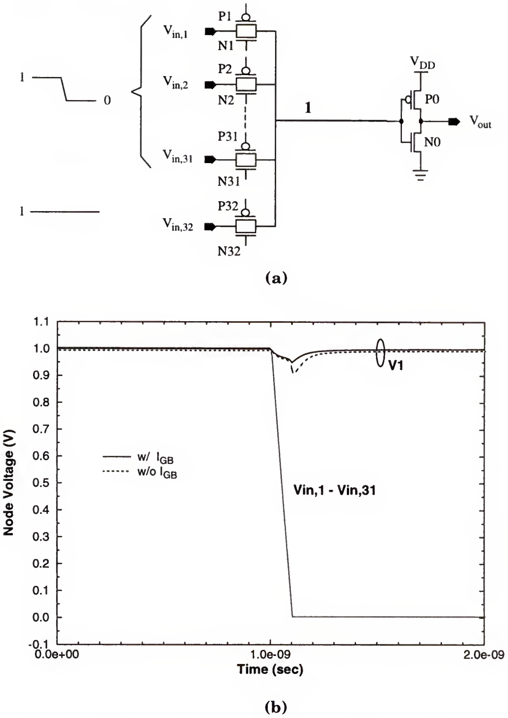


Figure 3.2 Pass transistor-based 32-to-1 multiplexer. (a) Schematic of the circuit. (b) The UFPDB/Spice3-predicted effect of the gate-to-body tunneling current on the node-1 voltage; L_{gate} = 60nm, t_{ox} = 1.4nm, and V_{DD} = 1.0V.

Figure 3.2(b) show that the node-1 voltage is not pulled down enough to upset the output logic state even if we consider the case without I_{GB}. This is because of the low transient BJT current due to low V_{DD} as discussed in section 3.3. Thus, the suppression of the transient BJT effect due to the I_{GB} is not relevant in this case.

3.4 Dynamic Logic

3.4.1 Dynamic 4-Way OR Circuit

For the pass transistor-based wide multiplexer in Figure 3.2, node-1, which is pulled down by transient BJT current, is held/restored by the single selected pass gate. The effect of the transient BJT current is only a very small dip in the node-1, and thus the I_{GB} effect is not relevant. For dynamic circuits [Wes93], the effect of the transient BJT current can be much more severe, and thus the I_{GB} effect can be important in the circuit. First, consider the dynamic four-way OR circuit in Figure 3.3(a), where N1/N2/N3/N4 are the logic transistors, and P0 and N0 are the clocked precharge and evaluation transistors, respectively, operating at V_{DD} = 1.0V. The feedback half-latch pMOSFET, which holds the dynamic node-2 [Lu97], is eliminated to consider the worst-case operation. Assume in a precharge phase that the input to N1 is at “1” and the inputs to N2/ N3/N4 are at “0” state. The dynamic node-2 is at V_{DD}, and the common source node-1 is at (V_{DD} - V_t). The input to N1 switches at t = 0.6ns from “1” to “0”, and the circuit subsequently evaluates at t = 1.1ns. Transient BJT currents flow through the off devices (N1/N2/N3/N4) as V1 drops to

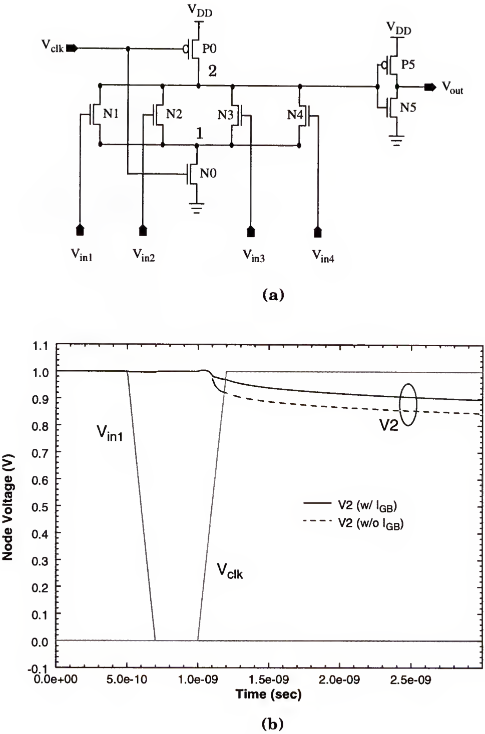


Figure 3.3 A dynamic four-way OR circuit. (a) Schematic of the circuit. (b) The UFPDB/Spice3-predicted effect of the gate-to-body tunneling current on the node-2 voltage. The width of each 60nm transistor (with t_{ox} = 1.4nm) is set to maintain the same W/L ratio given in Figure 3(a) of [Lu97]; V_{DD} = 1.0V.

"0", and they tend to pull down V2, and cause an erroneous evaluation. The UFPDB/Spice3 simulation results in Figure 3.3(b) show that the transient BJT currents produce a transient voltage dip in V2 to 0.85V without I_{GB}, while I_{GB} limits this dip to 0.90V. Because of the low transient BJT current due to low V_{DD}, no upset occurs (i.e., V2 is not pulled down enough to cross the threshold of the output inverter) in either case, but we note that I_{GB} does result in less dynamic power consumption and more noise margin.

3.4.2 Pseudo Two-Phase Dynamic 2-Way NAND Circuit

The transient BJT current can also potentially perturb the logic state in pseudo two-phase dynamic logic [Lu97] as shown in Figure 3.4(a). The configuration of pseudo two-phase dynamic logic is composed of two stages of dynamic logic whose inputs are controlled by the pass gate for each stage. The pass gates are controlled individually by two non-overlapping clocks *clk1* and *clk2*. While the first stage is evaluating by "0" of *clk1* ("1" of $\overline{clk1}$) as shown in Figure 3.4(a), the second stage is precharged by "1" of *clk2* ("0" of $\overline{clk2}$), and simultaneously the output of the first stage is stored in the gate capacitances of the second stage N3. The subsequent "0" of *clk2* ("1" of $\overline{clk2}$) forces the second stage to be evaluated, and the corresponding output is latched in a succeeding "1" of *clk1*.

The time period when *clk2* is at "0" is considered to examine transient BJT-current and corresponding I_{GB} effects. In this time period,

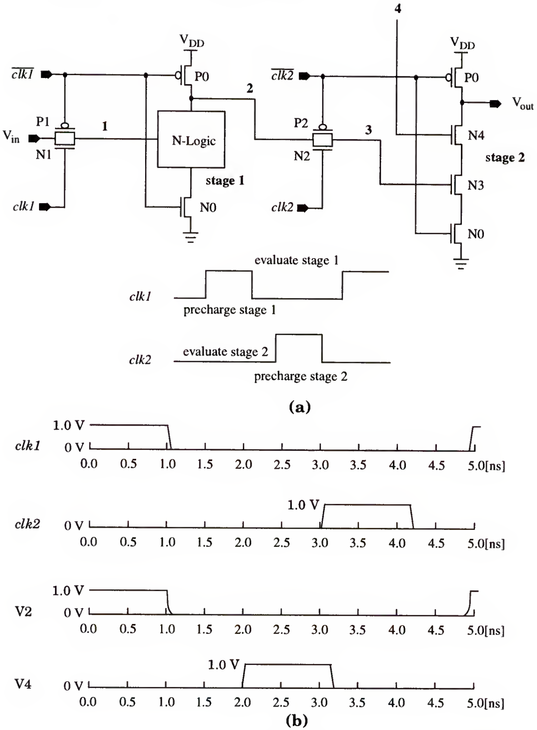


Figure 3.4 A pseudo two-phase dynamic 2-way circuit. (a) Schematic of the circuit. (b) The timing diagram; L_{gate} = 60nm, W_{gate} = 5µm (N0, N3, N4, P0) and 20µm (N1, N2, P1, P2), t_{ox} = 1.4nm, and V_{DD} = 1.0V.

the pass gate to the second stage is off and the second stage is evaluating. If the state stored in the gate capacitance of N3 is at "1" (node-3 is at "1") and the first stage is evaluated to be "0" (node-2 is at "0") at the falling edge of *clk1*, transient BJT current flows through the nMOSFET in the off-state pass gate, and the current discharges the input node (node-3) from "1" to "0" state. The transient BJT current is detrimental to the circuit operation via the discharge of the input node voltage for dynamic NAND logic. A possible example is illustrated in the timing diagram in Figure 3.4(b). During the evaluation phase for the second stage (*clk2* is at "0"), both inputs (node-3 and node-4) are at "1" to yield a "0" state for the output V_{out}. If the node-3, which is held by the gate capacitance of N3, is discharged completely by the transient BJT current at t = 1ns as in Figure 3.4(b), and the "1" signal arrives at node-4 at t = 2ns, the NAND logic can be erroneous as the output V_{out} remains in the logically wrong "1" state. The UFPDB/Spice3 simulation results in Figure 3.5 show that the node-3 voltage is brought down to 0.70V by the transient BJT currents without I_{GB}, while I_{GB} limits this to 0.80V. Because of the low transient BJT current due to low V_{DD}, no upset occurs in either case, but we note that I_{GB} does result in less dynamic power consumption and more noise margin.

3.4.3 Dynamic Cascade Voltage Switch Logic Circuit

Dynamic cascade voltage switch logic (dynamic CVSL) was investigated as another circuit topology susceptible to the transient BJT

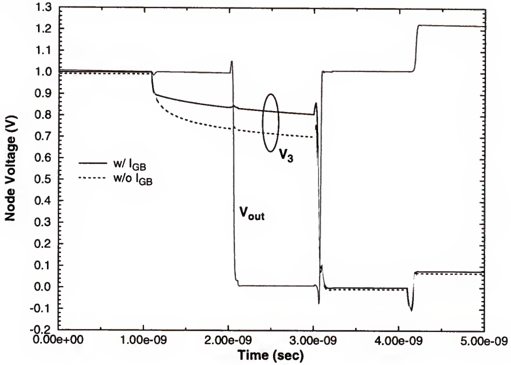


Figure 3.5 The UFPDB/Spice3-predicted effect of the gate-to-body tunneling current on the node-3 voltage: L_{gate} = 60nm, t_{ox} = 1.4nm, and V_{DD} = 1.0V.

effect [Chu97]. We show how the I_{GB} can influence the transient BJT effects in the circuit using UFPDB/Spice3. Figure 3.6 shows the schematics of a 3-input dynamic CVSL XOR circuit, where P0/P1 and N0 are the clocked precharge and evaluation transistors respectively. P11/ N11 and P12/N12 are the output inverters. The feedback half-latches to hold the dynamic node-1 and node-2 to improve the noise margin in Fig. 2 of [Chu97] are eliminated to consider worst-case operation. The width of each 60nm transistor is set to maintain the same W/L ratio given in Fig. 2(b) of [Chu97] as in Figure 3.6.

In the precharge phase (*clk* is at "0"), the clocked evaluation transistor N0 is off and the precharge transistors P0/P1 are on. For any input condition, all common-source nodes in all cascade levels are at "1" because every inputs are differential. Therefore, the logic transistors with "0" input are susceptible to transient BJT current because their drains and sources are at "1", and thus the bodies are at "1" if enough time is given to yield their steady-state values before the circuit goes to the evaluation mode.

In the evaluation phase (*clk* is at "1"), with the input pattern (A, B, C) = (0, 1, 1), the node-3 is pulled down by the clocked evaluation transistor N0. Node-5 is pulled down by N2, which is turned on ($\overline{A} = 1$). Transient BJT current flows through the N1, and node-4 is brought down (supposedly remains at "1") as shown in Figure 3.7. For the upper level of transistors, node-7 is pulled down by N4 (B = 1) and transient BJT current

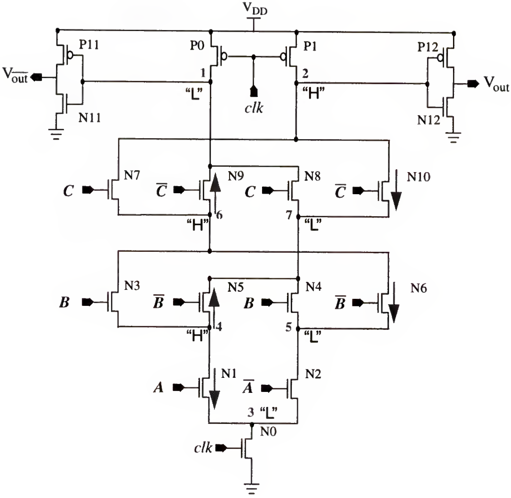


Figure 3.6 Schematic of 3-input dynamic CVSL XOR circuit; L_{gate} = 60nm, W_{gate} = 7.5μm (N0), 4.0μm (N1 - N10), 0.6μm (N11, N12), 5.0μm (P11, P12), and 1.2μm (P0, P1), t_{ox} = 1.4nm, and V_{DD} = 1.0V. "H"s and "L"s are the logic state for each node and the arrows represent the transient BJT current in the evaluation phase when the input condition is (A, B, C) = (0, 1, 1).

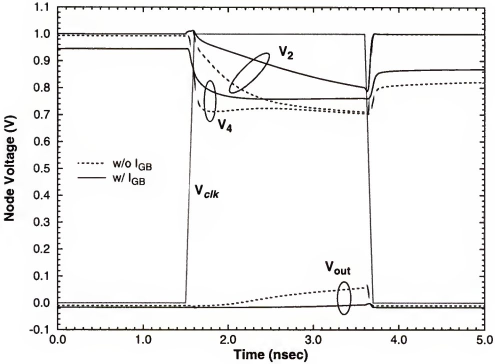


Figure 3.7 The UFPDB/Spice3-predicted effect of the gate-to-body tunneling current for dynamic CVSL XOR circuit: L_{gate} = 60nm, t_{ox} = 1.3nm, and V_{DD} = 1.0V.

flows through the N6 to pull-down node-6, which supposedly remains at "1" state. Note that pulling down node-7 by turning-on N4 induces another transient BJT current through N5, which enhances the transient voltage dip of node-4. The logic transistors of C - $\overline{C}$ level are operated by the same way as above. Node-1 is pulled down by N8 and transient BJT current flows through the N10 because node-7 goes to "0" state. The transient BJT current pulls down node-2, which supposedly remains at "1" as shown in Figure 3.7. Note again that pulling down node-1 by turning-on N8 induces another transient BJT current through N9. As can be seen in Figure 3.6, all the transient BJT currents contribute to pull down node-2 (supposedly remains at "1"). Logic error will not occur because the transient BJT current is not large enough to pull down node-2 across the threshold voltage of the output inverter (P12/N12). As can be clearly seen in Figure 3.7, however, I_{GB} does result in less dynamic power consumption and more noise margin via reducing transient BJT current.

3.5 Sense Amplifier

Finally, we consider a sense amplifier, in which the floating-body effects are typically severe and for which, as we show, I_{GB} can provide a benefit even in scaled PD SOI CMOS. Without I_{GB}, physical circuit simulations described in [Fos98], based on a 0.25μm CMOS technology, showed that dynamic instabilities (i.e., bit errors) result from threshold-voltage imbalances caused by hysteretic dynamic floating-body charging, necessitating body ties. Here, we examine the effect of the gate-to-body tunneling current on the dynamic instabilities in the same sense-amplifier circuit, but based on the contemporary L_{gate} = 60nm technology noted previously. The schematic of the full CMOS sense-amplifier circuit [Fos98] is shown in Figure 3.8. The width of each transistor is set proportionally to the corresponding W/L ratio used in the original design [Fos98]. The sense amplifier is composed of two coupled nMOS/pMOS pairs, N1/P1 and N2/P2. We assume that the storage capacitors, e.g., CS1 and CS2, are 15 fF, and the bitline capacitance, represented by C_{BL}, is 250 fF.

The normal sensing operation of the amplifier is indicated by the sequential pulses of V_{PRE}, V(WL), and V_{SE} as shown in Figure 3.8; we assume 10ns precharge and sense times with 0.5ns rise and fall times. The sense amplifier must respond to the small differential voltage (ΔV) established across the bitlines by activation of WL1 (while WL2 remains off). For example, when reading a "0" on CS1, $\Delta V < 0$ must result in N1 and P2 being turned on, and N2 and P1 off, which means V(BL) and $V(\overline{BL})$ go to 0 and V_{DD}, respectively. The bit is thus read, and the referenced cell is simultaneously refreshed. With regard to possible floating body-induced instabilities, N1 and N2, which drive the proper states of the coupled transistor pairs in response to ΔV, are the most crucial devices.

The UFPDB/Spice3 simulation results in Figure 3.9(a) show the body and source voltages of N1/N2 during a sequential sense cycle in the amplifier, after the circuit was held for a lengthy time with V(BL) = 0 and $V(\overline{BL}) = V_{DD}$ (as in an extended read- or write-"0"). In this state, V_{DS}(N2)

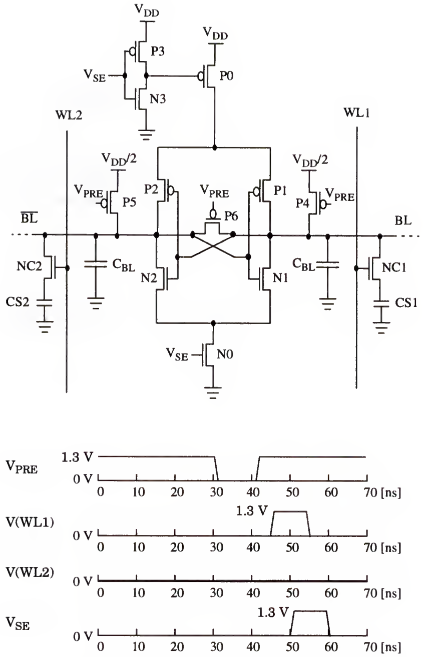
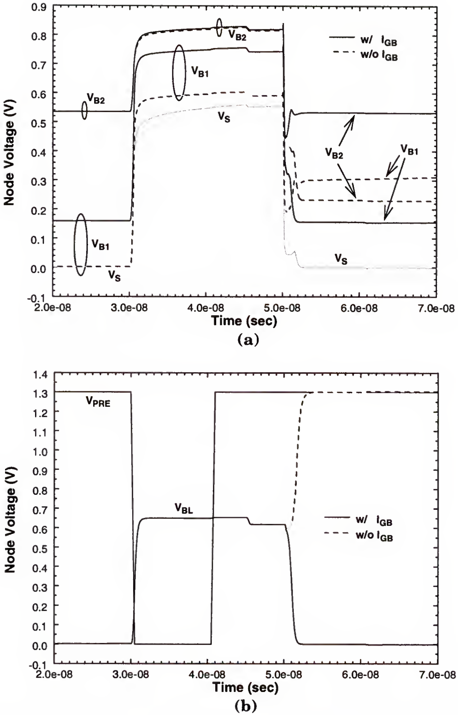


Figure 3.8 DRAM sense-amplifier circuit, including precharge and enable circuitry and showing two complementary data-storage cells on the bitlines and representative pulse sequences for sensing DRAM-cell data: $L_{gate} = 60nm$ and the width of each transistor is set proportionally to the corresponding W/L ratio used in the original design [Fos98].



(a)

(b)

Figure 3.9 UFPDB/Spice3 simulation results for the PD/SOI DRAM sense amplifier (L_{gate} = 60nm, t_{ox} = 1.4nm). (a) The predicted N1 and N2 body-voltage transients with the source voltage. (b) The bit-line voltage with and without gate-to-body tunneling current accounted for. For V_{DD} = 1.3V, as used here, and higher, I_{GB} prevents the erroneous sensing of a “1”; for lower V_{DD}, it does not.

$\cong V_{DD}$ and $V_{DS}(N1) \cong 0$, and hence $V_{BS}(N2) > V_{BS}(N1)$ as defined by the the body nodal equation ($I_R(V_{BS}) = I_G(V_{BD}) + I_{GB}(V_{GB})$) applied to both devices. The gate-to-body current (I_{GB}) of N2, which is conduction-band electron tunneling current for the depletion case (I_{CBE} in (2.26)), is negligible due to low negative V_{GB} (= $-V_{BS}(N2)$). However, I_{GB} of N1, which is valence-band electron tunneling current for the inversion case (I_{VBE} in (2.20)), is higher due to high V_{GS} (= V_{DD}), and can significantly reduce the steady-state and subsequent dynamic differences between $V_{BS}(N2)$ and $V_{BS}(N1)$ as shown in Figure 3.9(a). Hence, for high enough V_{DD}, as our simulation results in Figure 3.9(b) show, I_{GB} can actually prevent the bit error resulting from the floating-body instability noted in [Fos98].

The results in Figure 3.4 are explained as follows. In the precharge cycle, the drain node and source node of N1 are brought up from zero to $V_{DD}/2$ and $V_{DD}/2-V_t(t)$, respectively, since N1 is ON. The drain-body-source capacitive coupling in N1 causes the body voltage to increase by almost $V_{DD}/2$, while the gate-to-body capacitive coupling effect in N1 is negligible due to the shielding effect of the inversion layer. However, this is not the case for N2, where there is no inversion layer. The drain node in N2 is brought down to $V_{DD}/2$ and the gate and source nodes in N2 are brought up to $V_{DD}/2$ and $V_{DD}/2-V_t(t)$, respectively. The net capacitive coupling effect results in boosting $V_B(N2)$ up to 0.82V, irrespective of I_{GB}. However, I_{GB} in N1 makes $V_{BS}(N1)$ higher and thus decreases the imbalance of V_{BS} in N1 and N2. The subsequent dynamic $V_{BS}(N2) > V_{BS}(N1)$ imbalance is maintained when the amplifier is enabled at t = 50ns; $V_{DS}(N0)$ drops to 0, thereby increasing both $V_{BS}(N2)$ and $V_{BS}(N1)$ via the gate-to-body-source capacitive coupling, and the resulting $V_{BS}(t)$'s define unbalanced $V_t(t)$'s. This dynamic $V_t(N2) < V_t(N1)$ imbalance without I_{GB} is large enough to prevent a valid read-"0" operation, while the threshold voltage imbalance with I_{GB} is not enough to upset the operation, as can be seen in Figure 3.9(b). The presence of I_{GB} makes the coupled pairs in the sense amplifier flip properly and prevents the erroneous sensing of a "1" state.

All of the simulations discussed in this section did not account for any source/drain engineering to enhance carrier recombination and effect good control of the floating-body effects. In fact, as implied in Figure 3.9(a) by the high DC value of $V_{BS}(N2)$ in the sense amp, such control will be necessary in scaled PD/SOI to avoid high off-state currents without excessive increase in threshold voltage [Pel02]. Additional UFPDB/Spice3 simulation done with source/drain-junction recombination current increased (by more than an order of magnitude) show that all the detrimental floating-body effects are well controlled. Hence, we conclude that, in general, I_{GB} should be insignificant in dynamic operation of optimally scaled floating-body PD SOI CMOS circuits.

3.6 Conclusions

Using the I_{GB} model in UFPDB/Spice3, we have shown how the tunneling current tends to suppress dynamic floating-body effects in some PD/SOI circuits, providing benefits that follow from the complex interactions among the bias conditions, circuit topologies, and switching patterns. However, the significance of these benefits of I_{GB}, as well as its detrimental effects [Fun00, Jos01, Chu02, Poi02], tends to be diminished in dynamic operation of scaled PD SOI CMOS circuits because the supply voltage will be reduced and the scaled SOI technology will necessarily be optimized to suppress the detrimental DC floating-body effects on off-state current.

CHAPTER 4
PARASITIC BIPOLAR JUNCTION TRANSISTOR IN NONCLASSICAL ULTRA-THIN-BODY TRANSISTORS: MODELING, VERIFICATION, AND APPLICATION

4.1 Introduction

The parasitic BJT effects of the FB SOI MOSFET have constituted a critical issue for a long time because they induce a premature breakdown [You88, Cho91], latch [Che88, Cho91], transient BJT current [Pel95], and also amplify GIDL current [Che92]. While impact ionization-induced phenomena (premature breakdown and latch) have become mitigated due to scaled operation voltage, the transient BJT current and the amplified GIDL current are still major concerns for nonclassical UTB transistors regarding control of off-state leakage current in view of dynamic and static power consumption [Sak03, Fos99, Sch04].

However, the parasitic BJT model [Kri96] of UFDG [Fos03a], which is our process/physics-based compact model for UTB devices and circuits, is inadequate to describe the transient BJT action and the amplification of GIDL current in nonclassical UTB transistors because it does not correctly account for accumulation in undoped UTBs. In this chapter, the previous BJT model is modified and upgraded to correctly describe majority-carrier concentration in the body, which governs the parasitic BJT gain. Further, circuit application with the upgraded UFDG in Spice3 [Fos03a] is exemplified by simulations of a nonclassical dynamic logic circuit, which show the parasitic BJT effect on circuit operation and stress the need to account for it in IC design at future CMOS technology nodes.

4.2 Effects of Device Structure on Gate-Induced Drain Leakage Current

There are several components of off-state leakage current such as gate-to-body/channel tunneling, band-to-band tunneling between the body and drain, direct quantum tunneling from source to drain, drain-induced-barrier-lowering (DIBL)-enhanced subthreshold conduction, GIDL current, and gate direct tunneling between the gate and source/drain extension area in the gate overlap region. Among these leakage components, GIDL current can be the dominant one in nonclassical UTB transistors [Sch04], including the FDFET and the double-gate MOSFET (DGFET), for low-power applications because it can be enhanced by increased gate work function needed for threshold-voltage control, and it can be amplified by the inherent parasitic BJT.

For FDFETs and symmetrical DGFETs, midgap metal is needed for gate material, instead of heavily doped polysilicon, for threshold-voltage control, while, in asymmetrical DGFETs, counter-type polysilicon is needed for the back-gate material. Therefore, the gate work function of nonclassical UTB transistors will be larger than that of conventional MOSFETs, and thus the flat-band voltage of the gate-drain MOS structure will be higher, which implies higher GIDL current due to higher electric field at the gate-drain surface interface. With the assumption that the surface potential is pinned near $-E_g/q + V_{BD}$ when tunneling occurs, the normal surface field is expressed as [Kim01]

$$E_{sD} \cong \frac{V_{GD} - V_{FB}^{D} + \frac{E_g}{q} - V_{BD}}{3t_{oxf}} . \tag{4.1}$$

The higher flat-band voltage of the nonclassical UTB transistors makes the surface field higher (more negative), and then the GIDL current is enhanced via the following conventional model [Fos99]:

$$I_{GIDL} \cong W\Delta L_f \frac{A}{B}(-E_{sD})\exp\left(\frac{B}{E_{sD}}\right) \tag{4.2}$$

where A and B are functions of the bandgap and the carrier effective mass, and $W\Delta L_f$ defines the effective area over which GIDL occurs.

Figure 4.1 shows the dependence of GIDL current, derived from (4.1) and (4.2), on the work-function difference between the gate and drain-extension region. The exponential dependence of the GIDL current on E_{sD} makes the strong dependence of the current on gate-oxide thickness (t_{oxf}) as well as work-function difference. Note that the work-function differences when midgap metal and counter-type polysilicon are used as gate material induce huge GIDL currents. Figure 4.2 shows the effect of operation voltage ($V_{DD} = -V_{GD}$) on GIDL current for different gate materials. The weak dependence of the current on V_{DD} suggests that even if the voltage is scaled down, GIDL current would still be a dominant

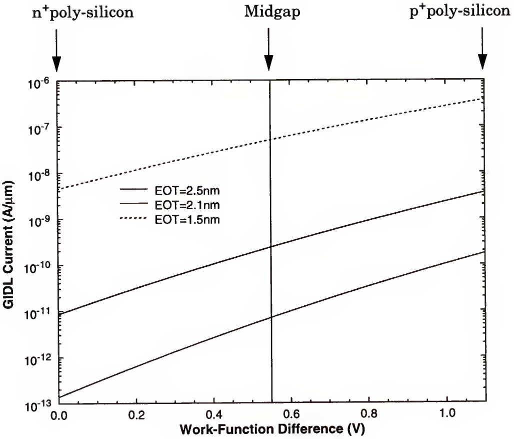


Figure 4.1 The predicted dependence of GIDL current on work function difference between gate and drain extension region in nMOSFET: V_{GD} = -1.2V, A = 2 × 10^9 A/cm^2, B = 4.5 × 10^9 V/m, and ΔL_f=10nm.

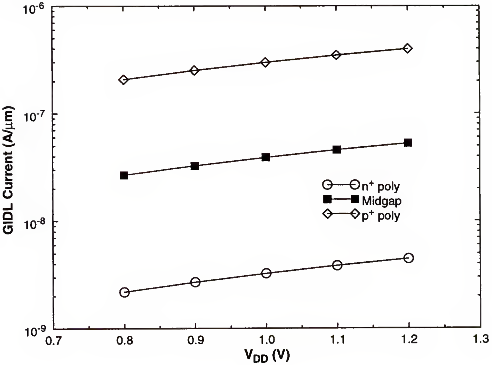


Figure 4.2 The predicted GIDL current vs. operation voltage: EOT = 1.5nm, A = 2×10^9 A/cm^2, B = 4.5×10^9 V/m, ΔL_f = 10nm, and $V_{GD} = -V_{DD}$.

component of off-state leakage current in nonclassical UTB transistors. Recently, a simulation-based study showed that a gate-underlap structure would be unavoidable and needed for nonclassical UTB transistors [Fos03b]. However, even for the gate-underlap structure, GIDL current, which can be induced by a gate fringing field, can still limit the off-currents for low-standby power devices [Sch04]. Therefore, reliable physical modeling of the parasitic BJT is needed to predict the amplification of GIDL current in nonclassical UTB transistors employing metal gates and undoped bodies.

4.3 Modeling of the Parasitic BJT

From the condition that the majority carriers are approximately in detailed balance, low-injection parasitic-BJT (base-transport) current density along the channel (y-) direction in a generic UTB nMOSFET, as illustrated in Figure 4.3, is expressed (in the forward-active mode) as [Kri96]

$$J_{nL}(x) = \frac{-qD_n n_i^2 \exp\left(\frac{V_{BS}}{V_T}\right)}{\int_0^{L_{eff}} p(x, y)dy} \tag{4.3}$$

where the suffix nL stands for electron and low injection, respectively; $V_T = kT/q$ is the thermal voltage. Note that V_{BS}, the body-source voltage, is defined in UFDG via the body nodal equation involving recombination-generation and charging currents. (More generally, the BJT current

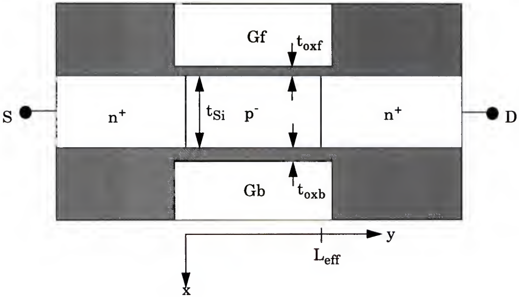


Figure 4.3 Schematic illustration of UTB nMOSFET, including the parasitic $n^+p^-n^+$ BJT.

includes an exp(V_{BD}/V_T) term [Kri96], as does our subsequent modeling of it.)

The denominator of (4.3) can be approximated under low-injection conditions to be [Kri96]

$$\int_0^{L_{eff}} p(x, y)dy \cong p(x)L_{eff} \tag{4.4}$$

where p(x) is the gate-controlled hole density in the UTB at depth x where the potential is minimum along y-direction; L_{eff} is the effective channel length assumed to be the BJT (quasi-neutral in y) base width. The approximation in (4.4) introduces some error, which can be tuned out via a model parameter, but most importantly conveys the bias dependence of the integrand. Using the Boltzmann approximation for carrier distribution, the hole density in the body can be expressed by

$$p(x) = N_b \exp\left[-\frac{\psi(x) - V_{BS}}{V_T}\right] \tag{4.5}$$

where N_b is the body doping density and ψ(x) is the potential at a point x in the body, referenced to that of a hypothetical, unbiased neutral body. Note that the front surface potential (ψ_{sf}) is assumed in the previous model [Kri96] to be V_{BS} in accumulation, and thus the hole density in accumulation is just the body doping density. This assumption is good for a high-doped body (as in PD/SOI MOSFETs in chapter 3). However, for undoped-UTB transistors, the hole density in accumulation can exceed N_b, and thus the previous BJT model largely overestimates the parasitic BJT current, as in [Fos99]. Therefore, the potential distribution in the body is needed for accumulation, as well as weak and strong inversion, to calculate p(x) and J_{nL} in (4.3).

For inversion, we use the solution of potential in the weak-inversion model [Yeh95] of UFDG to calculate the hole concentration in the body, and also use the same model for strong inversion, but pinning the potential to (ϕ_C ($\approx$ 0.4V) + ϕ_F) [Tri04] using an appropriate smoothing function. The loose validity of the potential model in strong inversion is mitigated by unnecessity of the BJT model in strong inversion where the BJT effects are insignificant for scaled CMOS technology due to scaled V_{DD}. The potential model in strong inversion is only needed for numerical stability.

In order to calculate the potential distribution in the body for accumulation, we relate the hole concentration to the accumulation charge in the body as

$$q\int_0^{\frac{t_{Si}}{2}} p(x)dx = Q_{accf} \tag{4.6}$$

$$q\int_{\frac{t_{Si}}{2}}^{t_{Si}} p(x)dx = Q_{accb} \tag{4.7}$$

where Q_{accf} and Q_{accb} are components of the integrated accumulation charge in the body, which is partitioned at the center as indicated. We assume Q_{accf} and Q_{accb} can be approximated as the respective gate charges:

$$Q_{accf} \cong -C_{oxf}(V_{GfS} - V_{BS} - \Phi^f{}_{MS} + \phi_C - \phi_F) \tag{4.8}$$

$$Q_{accb} \cong -C_{oxb}(V_{GbS} - V_{BS} - \Phi^{b}{}_{MS} + \phi_C - \phi_F)\,. \tag{4.9}$$

Analogous to the strong-inversion assumption, the surface potentials (ψ_{sf} in (4.8) and ψ_{sb} in (4.9)) are estimated to be pinned at (V_{BS} -ϕ_C + ϕ_F) for strong accumulation. These estimations are not critical because the significant dependences on the actual ψ_{sf} and ψ_{sb} will be conveyed by the left-hand sides of (4.6) and (4.7). The front and back accumulation charges in (4.8) and (4.9) are reduced to zero exponentially with increasing gate bias as in UFDG:

$$Q_{accf}' = C_{oxf} C \ln\left\{1 + \exp\left[-\frac{(V_{GfS} - \Phi^{f}{}_{MS} - V_{BS} + \phi_C - \phi_F)}{C}\right]\right\} \tag{4.10}$$

$$Q_{accb}' = C_{oxb} C \ln\left\{1 + \exp\left[-\frac{(V_{GbS} - \Phi^{b}{}_{MS} - V_{BS} + \phi_C - \phi_F)}{C}\right]\right\} \tag{4.11}$$

where C is constant which defines the stiffness of the smoothed charge transition between accumulation and inversion.

To get an analytical expression of (4.6) and (4.7), first for symmetrical DGFETs, the potential distribution in the body is approximated as linear in x between the surfaces and the center as in Figure 4.4:

$$\psi(x) = \psi_{sf} - \frac{\psi_{sf} - \psi_c}{t_{Si}/2}x \quad \text{for } 0 < x < t_{Si}/2 \tag{4.12}$$

$$\psi(x) = \psi_c - \frac{\psi_c - \psi_{sb}}{t_{Si}/2}x \quad \text{for } t_{Si}/2 < x < t_{Si}\,. \tag{4.13}$$

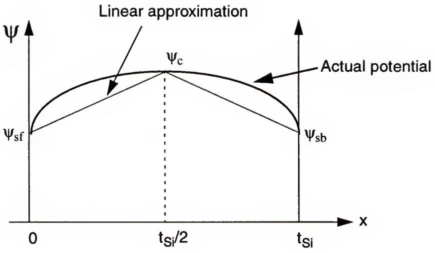


Figure 4.4 Potential, ψ(x), distribution in the vertical direction and the linear approximation for symmetrical double gate MOSFET.

Substituting (4.5) with (4.12) and (4.13) into (4.6) and (4.7) and carrying out the integration yields

$$\frac{1}{2}qN_b t_{Si}\left[\frac{\exp\left(-\frac{\psi_c - V_{BS}}{V_T}\right) - \exp\left(-\frac{\psi_{sf} - V_{BS}}{V_T}\right)}{\left(-\frac{\psi_c}{V_T}\right) - \left(-\frac{\psi_{sf}}{V_T}\right)}\right] = Q_{accf}' \tag{4.14}$$

$$\frac{1}{2}qN_b t_{Si}\left[\frac{\exp\left(-\frac{\psi_{sb} - V_{BS}}{V_T}\right) - \exp\left(-\frac{\psi_c - V_{BS}}{V_T}\right)}{\left(-\frac{\psi_{sb}}{V_T}\right) - \left(-\frac{\psi_c}{V_T}\right)}\right] = Q_{accb}' . \tag{4.15}$$

The potential ψ_c at the body center is estimated by adding αV_T, which is determined by numerical simulation and adjusted by a model parameter α, to the average value of the surface potentials:

$$\psi_c = \frac{\psi_{sf} + \psi_{sb}}{2} + \alpha V_T . \tag{4.16}$$

Using (4.16) and equating ψ_{sf} to ψ_{sb} for the symmetrical DG device then yields

$$\psi_{sf} = -V_T \ln\left[\frac{2\alpha Q_{accf}'}{qN_b t_{Si}(1 - e^{-\alpha})}\right] + V_{BS} = \psi_{sb} , \tag{4.17}$$

now with

$$\psi_c = \psi_{sf} + \alpha V_T . \tag{4.18}$$

Note that (4.17) and (4.18) analytically relate ψ to $V_{GfS} = V_{GbS}$.

For FDFETs with UTBs, the results of 2-D numerical device simulations with MEDICI [Med03] in Figure 4.5 suggest that ψ_c can be assumed to equal to ψ_{sb}, for varying t_{Si}, at least for long L_{eff}. The FD

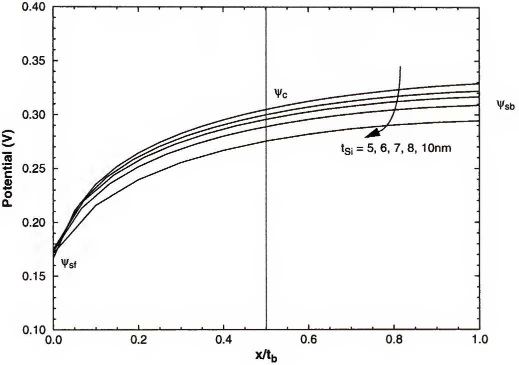


Figure 4.5 MEDICI-simulated potential distribution (center of channel: y = 50nm) of FD nMOSFET with varying t_{Si} along vertical direction (x-direction); L_{met} = 100nm, t_{ox} = 1.1nm, t_{BOX} = 200nm, midgap gate, $N_B = 1.0x10^{10}cm^{-3}$, V_{GfS} = -0.5V, V_{GbS} = 0.0V, and V_{DS} = 1.2V.

nMOSFET simulated has a long channel (L_{met} = 100nm), with t_{BOX} = 200nm, t_{ox} = 1.1nm, and undoped body ($N_b = 1\times10^{10}\text{cm}^{-3}$). The simulation results shown are the potential ψ(x) at the center of channel, y = $L_{eff}/2$. Carrying out the integration in (4.7) with this assumption gives

$$\psi_{sb} = -V_T\ln\left[\frac{2Q_{accb}'}{qN_bt_{Si}}\right] + V_{BS} = \psi_C. \tag{4.19}$$

Now, to get an analytical expression for ψ_{sf} for the FDFET, the potential distribution is approximated as constant in x between the front surface and the center as illustrated in Figure 4.6. This approximation ultimately yields a high ψ_{sf}, and the BJT current is overpredicted at the front surface. However, the parasitic BJT current is predominant between the center of the body and the back surface due to relatively low p(x) there. Thus, the error due to the approximation in Figure 4.6 is negligible for the common FDFET structure, which has a thick buried oxide. Carrying out the integration in (4.6) then gives

$$\psi_{sf} = -V_T\ln\left[\frac{2Q_{accf}'}{qN_bt_{Si}}\right] + V_{BS}. \tag{4.20}$$

Because of the uncertainty of the charge model in (4.6) - (4.9) for the FDFET, a model tuning parameter (α) is used to adjust (4.19):

$$\psi_{sb}'(=\psi_c') = \psi_{sb}(=\psi_c) - \alpha V_T\ . \tag{4.21}$$

Hence, ψ is analytically related to V_{GfS} and V_{GbS} for the FDFET.

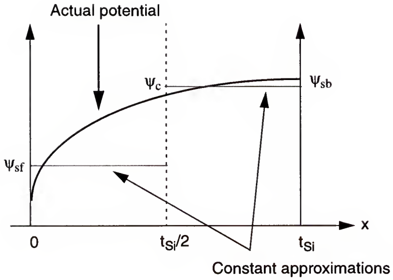


Figure 4.6 Potential, $\psi(x)$, distribution in the vertical direction and the constant approximations for FDFET

Now, for both the DGFET and the FDFET, the potential in the body for all bias regions is obtained by linking the potential model for accumulation to that for inversion by an appropriate smoothing function:

$$\psi_t = \psi_a + \frac{\ln\{1 + \exp[C(\psi_{ws} - \psi_a)]\}}{C} . \tag{4.22}$$

where ψ_a represents ψ_{sf}, ψ_{sb}, or ψ_c for accumulation calculated by (4.16) - (4.21), and ψ_{ws} represents each potential in weak inversion, which is derived from the weak-inversion solution [Yeh95] of UFDG. Using the solution of each potential (ψ_t in (4.22)) in the body, the low-injection BJT current is completely described by (4.3) - (4.5).

The high-injection BJT current density is written from standard analysis [Mul85],

$$J_{nH} = \frac{-2qD_n n_i \exp\left(\frac{V_{BS}}{2V_T}\right)}{L_{eff}}, \tag{4.23}$$

and the net current density for arbitrary injection level is given as [Kri96]

$$J_{nT}(x) \cong \frac{J_{nL}(x) J_{nH}}{J_{nL}(x) + J_H} . \tag{4.24}$$

Substituting (4.3), (4.4), and (4.23) in (4.24) yields

$$J_{nT}(x) \cong \frac{-2qD_n n_i^2 \exp\left(\frac{V_{BS}}{V_T}\right)}{L_{eff}} \left[\frac{1}{2p(x) + n_i \exp\left(\frac{V_{BS}}{2V_T}\right)}\right] . \tag{4.25}$$

The total BJT current is calculated by integrating (4.25) between x = 0 and x = t_{Si}:

$$I_{BJT} = -W\int_0^{t_{Si}} J_{nT}(x)dx \tag{4.26}$$

where W is the device width. To evaluate (4.26), a linear variation of potential in the body is assumed again as shown in Figure 4.4. Using (4.5), (4.12), and (4.13), and carrying out the integration in (4.26) with (4.25) yields

$$I_{BJT} = W\frac{qD_n n_i^2}{L_{eff}}\frac{1}{p_H}\{F_1 + F_2\}\exp\left(\frac{V_{BS}}{V_T}\right) \tag{4.27}$$

with

$$F_1 = \frac{1}{a_f}\ln\left[\frac{(t3-1)t1}{(t1-1)t3}\right] \tag{4.28}$$

$$F_2 = \frac{1}{b_f}\ln\left[\frac{(t2-1)t3}{(t3-1)t2}\right], \tag{4.29}$$

where $p_H = n_i\exp\left(\frac{V_{BS}}{2V_T}\right)$, $a_f = \frac{\psi_{sf} - \psi_c}{V_T t_b}$, $b_f = \frac{\psi_c - \psi_{sb}}{V_T t_b}$, t1 = 1 + 2p(x = 0)/$p_H$, t2 = 1 + 2p(x = t_{Si})/p_H, and t3 = 1 + 2p(x = t_{Si}/2)/p_H. When ψ_{sf} equals ψ_c or ψ_c equals ψ_{sb}, (4.28) and (4.29) are invalid and induce numerical overflow. In that case, F_1 and F_2 in (4.28) and (4.29) are smoothly changed to t_{Si}/t_1 and t_{Si}/t_2 independently. Eqation (4.27) defines the BJT current for forward mode. In the reverse mode, the current is expressed as in (4.27), with V_{BD} replacing V_{BS}. The composite parasitic BJT current is then given by the summation of the forward-mode and reverse-mode components.

For the transient charge dynamics, the bipolar charges associated with each of the terminals must be properly accounted for as in [Kri96]. Of the two main components of stored bipolar charge (minority carriers stored in the body and in the source), the previous model of minority carriers injected in the body should be modified because it depends on the majority-carrier distribution in the body, which is modified by the new modeling for accumulation. The minority-carrier charge injected in the (p-type) body is given by

$$Q_n = -Wq\int_0^{t_{Si}}\int_0^{L_{eff}} \Delta n(x, y)dxdy. \tag{4.30}$$

Assuming a linear distribution of electrons along the channel/body (base), for all injection levels, the excess electron density can be expressed as

$$\Delta n(x, y) \cong \Delta n(x)\left(1 - \frac{y}{L_{eff}}\right). \tag{4.31}$$

Substituting (4.31) in (4.30) gives

$$Q_n \cong -Wq\frac{L_{eff}}{2}\int_0^{t_{Si}} \Delta n(x)dx. \tag{4.32}$$

The composite Δn(x) is given by a weighted sum analogous to (4.24):

$$\Delta n(x) \cong \frac{\Delta n_L(x)n_H}{\Delta n_L(x) + n_H} \tag{4.33}$$

where

$$\Delta n_L(x) \cong \frac{n_i^2}{p(x)}\left[\exp\left(\frac{V_{BS}}{V_T}\right) - 1\right] \tag{4.34}$$

and $n_H = n_i\exp(V_{BS} / 2V_T)$. Using (4.33) and (4.34) in (4.32) and carrying out the integration yields the total charge stored in the body as

$$Q_n \cong -\frac{WqL_{eff}}{4}\left[\frac{1}{a_f}\log\left\{\left(\frac{s2-f1}{s1-f1}\right)\left(\frac{s1}{s2}\right)\right\}+\frac{1}{b_f}\log\left\{\left(\frac{s4-f1}{s3-f1}\right)\left(\frac{s3}{s4}\right)\right\}\right]p_H \qquad (4.35)$$

where $f1 = \exp(V_{BS} / V_T)$, $f2 = p(x = 0) \exp(V_{BS} / 2V_T) / n_i$, $f3 = p(x = t_b/2) \exp(V_{BS} / 2V_T) / n_i$, $S1 = f1 + f2$, $S2 = f1 + f2 \exp(a_f t_b)$, $S3 = f1 + f3$, and $S4 = f1 + f3 \exp(b_f t_b)$. For the reverse mode, the stored charge is expressed as in (4.35) with V_{BD} replacing V_{BS} in p_H, f1, f2, and f3.

4.4 Model Implementation/Verification

The upgraded model for the parasitic BJT in UTB devices was implemented in the UFDG model in Spice3. An existing model parameter, FVBJT, which was previously used to empirically account for breakdown in strong inversion, was removed since the parasitic BJT is no longer important in strong inversion due to scaled operation voltage. A new model parameter FABJT (α in (4.18) for the symmetrical DGFET and in (4.21) for the FDFET) was added for tuning the BJT gain in the accumulation bias condition. Note that the new BJT model is physically valid only for symmetrical DG and FD MOSFETs. For asymmetrical DG devices, it is largely empirical.

To predict the parasitic BJT gain of nonclassical UTB MOSFETs reliably, the calibration of the smoothing-function constant C of (4.22) is critical because it governs the body potential for weak accumulation, which is a common bias condition related to GIDL current and the transient BJT current of a pass gate. The smoothing-function constant is calibrated via numerical simulation of the transient BJT current using MEDICI. Table 4.1 gives structural information for the simulated undoped-UTB devices. L_{met} (= L_{eff}) was defined where the source/drain doping density drops to 10^{15} cm^{-3}. For the transient BJT-current simulation, the default Auger recombination and band-gap narrowing models in MEDICI were used. The UFDG model parameters related to device structure were directly adapted from the MEDICI-defined structure. (Note that quantization effects are negligible in these undoped devices for weak-inversion/accumulation conditions.)

Table 4.1: Device structure information

Device parameter	DGFET	FDFET
Gate length (L_{gate})	48.0 [nm]	48.0 [nm]
Metallurgical gate length (L_{met})	28.0 [nm]	28.0 [nm]
Front oxide thickness (t_{oxf})	2.0 [nm]	1.1 [nm]
Back oxide thickness (t_{oxb})	2.0 [nm]	200.0 [nm]
Silicon film thickness (t_{Si})	8.0 - 14.0 [nm]	5.0 - 8.0 [nm]
Body doping density (N_b)	Intrinsic	Intrinsic
Source/Drain doping density (N_{SD})	10^{20} [cm^{-3}]	10^{20} [cm^{-3}]
Source/Drain doping gradient	1.0 [nm/dec.]	1.0 [nm/dec.]
Gate work function (Φ_G)	Midgap + 0.1 [V]	Midgap + 0.1 [V]

Figure 4.7 shows UFDG-predicted transient BJT current characteristics of the symmetrical DGFET with different UTB thickness, compared with MEDICI predictions for the pass-transistor circuit. Note that the dependence of parasitic BJT gain on t_{Si}, reflected by the transient current (which was almost 20 times larger in the previous model for the same device structure) is modeled well. The threshold voltage of the device with thicker t_{Si} is lowered by SCEs, and hence the hole density in the body, which defines the transient BJT current, is lowered. Thus, the symmetrical DGFET having thicker t_{Si} shows higher BJT gain, and the upgraded BJT model predicts it well.

The same model calibrated to the DGFET is applied to the FDFET with device parameters in Table 4.1 to further verify the new BJT model. The model-predicted transient BJT current is in good agreement with the results of MEDICI simulations as shown in Figure 4.8. The smoothing-function constant, $C = 7.0$ for all DG and FD devices, is thus well defined. For the calibrations, α in (4.18) for the symmetrical DGFET and in (4.21) for the FDFET was 1.5 and 3.5, respectively, which are reasonable values. The physical nature of the new BJT model is thus confirmed.

The BJT model with C evaluated via the transient BJT current from MEDICI simulations, is next applied to GIDL current amplification. The measured GIDL data of UTB FDFETs [Kri02], which have a midgap metal gate and undoped body, are used. Calibrating UFDG [Fos03a] to

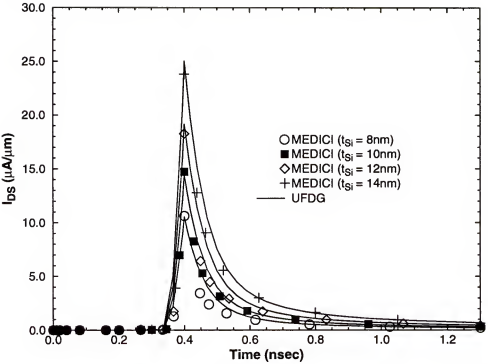


Figure 4.7 UFDG- and MEDICI-predicted transient BJT current (see Figure 3.1) for symmetrical DGFETs having several UTB thicknesses. The supply voltage V_{DD} was 1.2V and the fall rate of the source-voltage pulse was 1.2V/0.2ns. Note that I_{DS} is normalized with respect to the physical width of the DG devices.

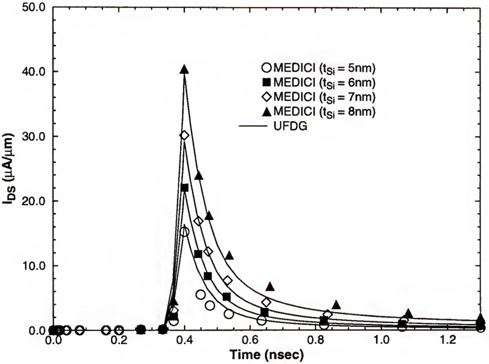
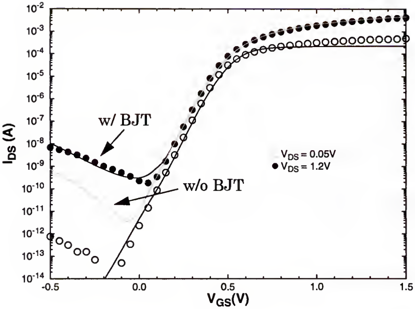


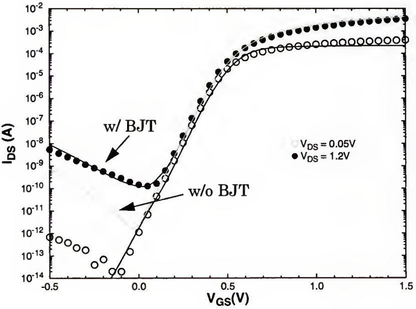
Figure 4.8 UFDG- and MEDICI-predicted transient BJT current for FDFET having several UTB thicknesses. The supply voltage V_{DD} was 1.2V and the fall rate of the source-voltage pulse was 1.2V/0.2ns.

measured I_{DS}-V_{GS} data, and tuning α to 2.5, we model the GIDL current, amplified by the BJT, to be quite consistent with the measured data for two different channel lengths, as shown in Figure 4.9. Hence, the BJT model verification is solidified.

The parasitic BJT gain implied by GIDL current also depends on silicon film thickness, as reflected by the transient BJT current in Figures 4.7 and 4.8. The validity of the BJT model for GIDL current versus t_{Si} is examined by MEDICI simulation. Due to the uncertainty of the band-to-band tunneling model in MEDICI, only the BJT gain (β), calculated by dividing the net amplified drain current by the GIDL current (which is obtained from the total band-to-band tunneling current in the output file of MEDICI), is compared with that calculated from UFDG simulation. Figure 4.10 shows the $\beta(t_{Si})$ results of UFDG and MEDICI simulations for FDFETs and symmetrical DGFETs with L_{eff} = 28nm at two different gate biases. The UFDG results show good agreement with the simulation results of MEDICI at weak accumulation (V_{GS} = -0.15V) for both devices, except for the thickest t_{Si}. This discrepancy is related to the low threshold voltage implied by worse SCEs. For these thick devices, the channel current is comparable to the GIDL-driven BJT current. Thus, the parasitic BJT gain inferred from MEDICI is overpredicted since channel current and the parasitic BJT current cannot be distinguished in the MEDICI-predicted drain current. This discrepancy thus does not necessarily reflect model error. The model does, however, show some error in strong

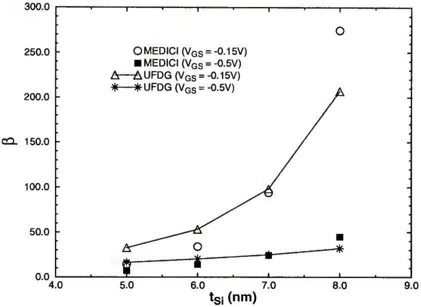


(a)

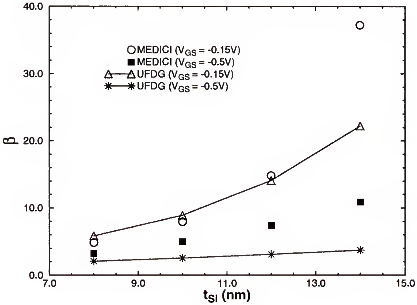


(b)

Figure 4.9 Measured [Kri02] (points) and UFDG--predicted current-voltage characteristics of a UTB FD nMOSFET showing the parasitic BJT amplification. (a) L_{gate} = 70nm, t_{oxf} = 2.1nm, t_{BOX} = 200nm, midgap gate, and undoped body. (b) L_{gate} = 80nm and the other parameters are same as in (a). The tuned value of B_{GIDL} is 3.8×10^9V/m.



(a)



(b)

Figure 4.10 UFDG- and MEDICI-predicted parasitic BJT gain for GIDL current. (a) FDFET. (b) Symmetrical DGFET.

accumulation (V_{GS} = -0.5V) for both devices. But, the strong-accumulation bias condition is not common in typical operation of digital logic circuits, and thus this deficiency of the model can be tolerable. Hence, overall, the parasitic BJT model is effective for the prediction of the dependences of BJT gain on reasonable bias conditions as well as on t_{Si} and L_{eff}.

Note that the BJT gain of the FDFET is much higher than that of the DGFET for the same t_{Si} in Figure 4.10. For the FDFET, the low hole concentration at the back surface induces large parasitic BJT current. However, t_{Si} of the FDFET needed for SCE control is thinner than that of the DGFET for the same L_{eff}, and thus the BJT gain of both devices with controlled SCEs will not differ substantively.

4.5 Model Application

To examine the parasitic BJT effects on circuit operation for nonclassical UTB devices, we simulate the dynamic 4-way OR circuit in Figure 3.3 of chapter 3, which can be affected by the transient BJT current, using UFDG in Spice3. The schematic of the circuit is redrawn in Figure 4.11. The DGFET for the simulation has L_{eff} = 28nm, $t_{oxf} = t_{oxb}$ = 2.0nm, intrinsic body, t_{Si} = 8nm, and midgap gate. For the FDFET, L_{eff} = 28nm, t_{oxf} = 1.1nm, t_{oxb} = 200nm, t_{Si} = 5nm, intrinsic body, and midgap gate are used. The width of each transistor is set proportionally to the corresponding W/L ratio used in the original design [Lu97]. As assumed in chapter 3, the dynamic node-2 is at V_{DD}, and the common source node-1 is at V_{DD}-V_t in precharge phase. When the circuit goes to the evaluation

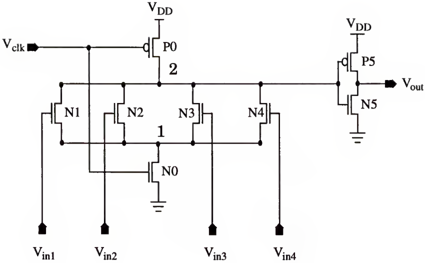
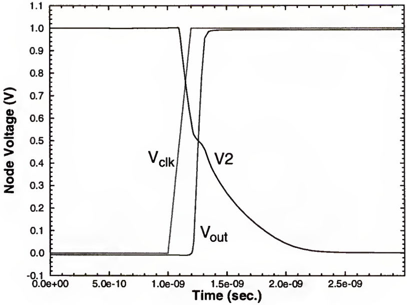


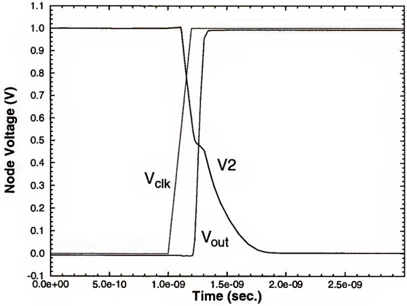
Figure 4.11 Schematic of a dynamic four-way OR circuit. The width of each 28nm transistor is set to maintain the same W/L ratio given in Figure 3(a) of [Lu97].

mode with “0” input for all logic devices (N1/N2/N3/N4), transient BJT currents flow through the all off devices (N1/N2/N3/N4) as V1 drops to “0”, and they pull down V2. The UFDG simulation results for both DGFETs and FDFETs in Figure 4.12 show that V2 is pulled down by the transient BJT current enough to cross the threshold of the output inverter, causing an erroneous evaluation. On the contrary, the transient BJT currents in the FB PD SOI circuit were not enough to upset valid operation due to low operation voltage in chapter 3 (see Figure 3.3(b)).

Similar to the results in Figure 4.7 and 4.8, the UFDG simulations of the pass-gate transistor here show that the transient BJT currents in the DGFET and the FDFET are comparable, and of the same order of magnitude as that in the FB PD SOI device for the same V_{DD} (= 1.0V). Unlike in classical CMOS, however, the DGFET and FDFET have very low intrinsic gate capacitance for low V_{GS} ($< V_t$) [Fos02b, Fos04]. For the evaluation mode of the circuit, the node-2 voltage is mainly held by gate capacitance of P5 and N5. Since P5 is in the off-state ($V_{GS} < V_t$) for the evaluation mode and the output inverter is skewed to speed-up the output pull-up transition (the width of P5 is 8 times larger than that of N5), the total capacitance to hold the node-2 voltage is much smaller compared to that of the FB PD SOI circuit (the gate capacitance of P5 is substantive in the off-state for the PD SOI device [Fos02b]). Thus, the small transient BJT current can pull down the node-2 voltage easily for the UTB devices. The feedback half-latch is needed to hold the dynamic



(a)



(b)

Figure 4.12 The UFDG predicted effect of the transient BJT current on the output voltage: $V_{DD} = 1.0V$. (a) Symmetrical DGFET. (b) FDFET.

node-2 [Lu97], and the size of the half-latch should be determined carefully to prevent logic upset and reduce noise margin via a reliable parasitic-BJT model for the UTB devices.

4.6 Conclusions

The effects of nonclassical undoped UTB device structures on GIDL current have been examined. GIDL current is enhanced by the needed higher gate work function of the UTB devices and amplified by the parasitic BJT. To predict the amplification of the GIDL current as well as transient BJT current, the parasitic BJT model for nonclassical UTB transistors has been upgraded and implemented in the UFDG MOSFET model. The model is verified with the results of MEDICI simulations and experimental data of FDFETs and symmetrical DGFETs, and shows good agreement for various device dimensions and reasonable bias conditions. For a dynamic logic circuit, we showed that the parasitic BJT can cause an erroneous evaluation for some bias conditions and switching patterns due to inherent low gate capacitance. Thus, a rigorous simulation is needed to ensure right operation and reduce noise margin for nonclassical UTB transistor circuits via a reliable parasitic BJT model for the UTB devices.

CHAPTER 5
TRIPLE-GATE MOSFET: INSIGHT, FEASIBILITY, AND ANALYSIS

5.1 Introduction

Conventional scaling of classical bulk-silicon and PD SOI CMOS has become very challenging because control of SCEs for acceptable I_{on}/I_{off} ratios requires precise channel doping levels and gradients [Fra01] that are, ultimately, impossible to achieve. Promising alternatives for continued CMOS scaling are the nonclassical devices having UTBs, i.e., the planar FDFET [Dor02] and the quasi-planar DG FinFET [Ked02], which give good SCE control via the UTB whether it is doped or not. The undoped UTB is the proper choice since it avoids the doping challenges noted for the classical devices, and the implied threshold voltage (V_t) variations [Xio03]. Further, it avoids source/drain junction-tunneling leakage currents and can perhaps yield high carrier mobilities due to negligible impurity scattering.

However, the good SCE control in the nonclassical nanoscale devices requires that the UTB thickness (t_{Si}) be considerably thinner than the gate length (L_g), as well as being uniform across the device and circuit. For the FDFET shown in Figure 5.1(a), t_{Si}, which is the height (h_{Si}) of the silicon film, should be scaled down to ~$L_{eff}/5$ [Tri03]. For the DG FinFET shown in Figure 5.1(b), t_{Si}, which is the width (w_{Si}) of the silicon fin,

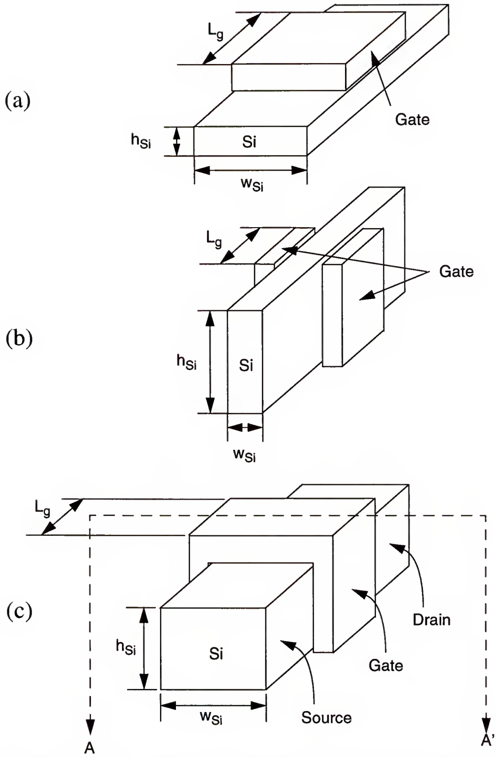


Figure 5.1 Schematic illustrations of nonclassical devices. (a) The FDFET. (b) The DG FinFET. (c) The TGFET. All devices are typically fabricated in SOI material, with the buried oxide (BOX) providing isolation.

should be scaled down to ~$L_{eff}/2$ [Fos03b], or about the minimum feature size in conventional CMOS technology. (Note that the effective channel length L_{eff} of nonclassical devices with undoped bodies will probably be longer than L_g [Fos03b].)

The TGFET was recently proposed [Doy03a, Doy03b] as a means to alleviate the stringent thin-t_{Si} requirements of UTB FDFETs and DG FinFETs, enabling "flexible and relaxed" silicon-body dimensions as illustrated in Figure 5.1(c). With lower h_{Si} and wider w_{Si}, the planar-like TGFET is easier to fabricate than the DG FinFET, and the experimental and simulation results in [Doy03a] and [Doy03b] suggest that excellent short-channel device performance can be achieved. These results, however, were obtained for TGFETs with heavily doped bodies, which, as noted above, are not desirable. Further, design studies of TGFETs [Doy03b] have revealed that the top corners of the silicon body (Figure 5.2) can, with strong dependence on their shape, or radius of curvature, significantly affect the channel current-voltage characteristics of the TG device. Typically, the corner components of current reflect a lower threshold voltage (V_t) and higher off-state current (I_{off}) relative to the respective properties of the bulk TG device, and for some body structures they can cause a "hump" in the subthreshold $\log(I_D)$-V_G characteristic [Doy03b]. Hence, a serious technological issue is implied. And, the layout area of the TGFET is an issue for nanoscale CMOS [Doy03b]. The feasibility of the TGFET is thus not yet clear.

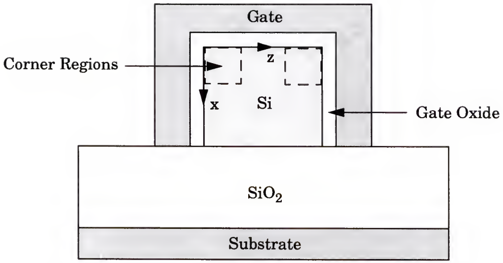


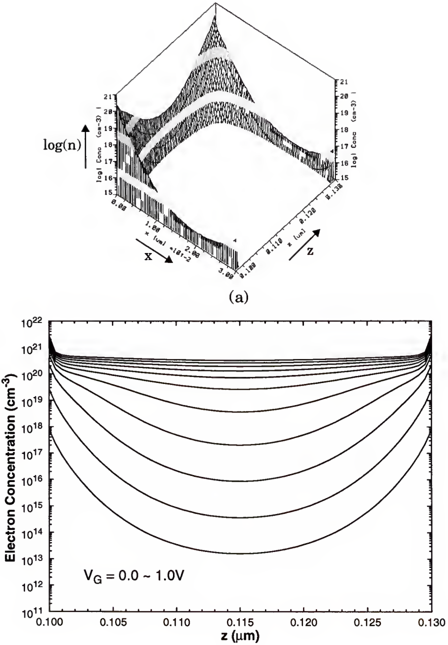
Figure 5.2 2D cross-sectional view A-A' of the TGFET, as indicated in Figure 5.1(c), identifying the critical corner regions of the silicon body/channel.

In this chapter, we examine the abnormal corner effects in nanoscale TGFET, using two-dimensional (2-D) numerical device simulations and quasi-2-D analysis to gain physical insights on how the effects could possibly be suppressed, irrespective of the radius of curvature, which is difficult to control technologically. An interesting and effective design regarding body/channel doping is proposed and explained. Then, we examine the feasibility of the TGFET for nanoscale CMOS applications with regard to SCEs and layout area. We check the SCEs for varying dimensions of the body, with and without doping, using Davinci [Dav03], a three-dimensional (3-D) numerical device simulator. Then, with the SCEs adequately controlled via body scaling in accord with the 3-D simulation results, we analyze the gate layout area of the integrated TGFET needed for current drive, and compare it with those of the FDFET and DG FinFET. The results show that the TGFET is not feasible, and that the DG FinFET, which is more scalable than the FDFET, has the most potential for future nanoscale CMOS applications.

5.2 Suppression of Corner Effects in TGFETs

Scaled TGFETs are 3-D devices (Figure 5.1(c)), and hence their complete characterization requires 3-D analysis, e.g., via computationally intensive 3-D numerical simulation as in [Doy03b]. However, to study the basic corner effects, we can analyze the 2-D cross section of the device, between the source and drain as shown in Figure 5.2, which is tantamount to assuming long gate length (L_g). We assume a rectangular SOI body, with height and width defined as h_{Si} and w_{Si}, respectively, sitting on the underlying buried SiO_2 (BOX)/Si substrate. A reasonable TGFET design, with a fully depleted body, could be $h_{Si} = w_{Si} = L_g$ [Cha02, Doy03b]. So, we first assume a scaled TG nMOSFET with $h_{Si} = w_{Si} = 30nm$, ignoring the SCEs, however, as noted above. We let the gate oxide thickness (t_{ox} = EOT) be 1.1nm, and the BOX thickness t_{BOX} = 200nm. Following [Doy03b], we further assume an n^+ polysilicon gate, and heavy doping in the (p-type) body for V_t control: $N_B = 8.0x10^{18}cm^{-3}$.

Modeling the substrate as simply a metal contact, which is grounded, we use MEDICI [Med99] to solve Poisson's equation in the defined 2-D SOI structure (Figure 5.2) for various gate voltages, V_G. (For thick t_{BOX}, effects of electric-field fringing and induced charge in the actual silicon substrate are small and can be ignored here.) Note that V_G equals the gate-source voltage (V_{GS}) of the TG nMOSFET when the source is grounded. A predicted, low-V_G 2-D inversion-electron distribution n(x, z) in the body cross-section is shown in Figure 5.3(a), and variations of n(x=0, z) across the top surface (channel) of the body are plotted, versus V_G, in Figure 5.3(b). Note, for the lower values of V_G which correspond to weak inversion, the dramatic build-ups of the electron density at and near the top corners (z = 0, W_{Si}) of the silicon body. They reflect the reduced V_t of the components of current there as revealed in [Doy03b]. In Figure 5.3(b), n(0, z=15nm) versus V_G reflects a bulk V_t of 0.3-0.4V, whereas the corner V_t is significantly lower. (Although we are not accounting for



(a)

(b)

Figure 5.3 MEDICI-predicted electron concentration of a high-doped TG nMOSFET. (a) 2-D electron distribution n(x, z'), where z' = z+100nm, in the silicon-body/channel; $h_{Si} = w_{Si} = 30$nm, $V_G = 0.3$V. (b) Predicted electron density n(x=0, z) across the top surface of the cross section for various values of V_G. Note, for this device with high $N_B = 8.0 \times 10^{18}$ cm^{-3}, the extremely high n in the corner regions for low V_G (subthreshold region). This electron buildup for high V_G is much less significant.

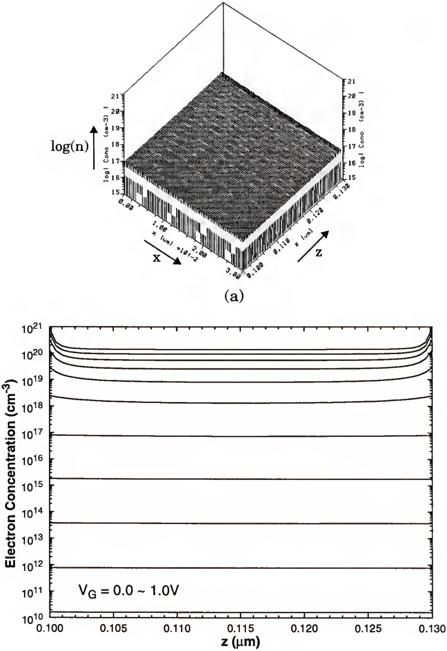
quantization (QM) effects [Ge02] here, our results concerning the corner effects are qualitatively valid. The QM solution, which requires a 2-D Poisson-Schrödinger solver that includes anisotropic effective mass modeling, would show, for a specific V_G, smaller n with the peak densities displaced from the surfaces, but still with relatively high densities in the corner regions.)

We can explain the inversion-charge build-ups at the corners by considering the 2-D Poisson equation for weak inversion ($n \ll N_B$) in the p-type silicon body (cross-section in Figure 5.2) of the TG nMOSFET:

$$\frac{\partial E_x}{\partial x} = -\frac{qN_B}{\varepsilon_{Si}} - \frac{\partial E_z}{\partial z} = -\frac{qN_{B(eff)}}{\varepsilon_{Si}} \tag{5.1}$$

where E_x and E_z are components of the electric field. In the vicinity of the left-side corner ($z \to 0$), the signs of both partial derivatives in (5.1) are negative for a given $V_G > V_{FB}$ (as in the respective 1-D versions of the equation). Thus, the effect of the left gate (i.e., E_z) is to reduce the magnitude of the E_x partial derivative, or of the spatially dependent effective doping density ($N_{B(eff)}$) as defined in (5.1). This reduction of $N_{B(eff)}$, which is enhanced for $z \to 0$ (and for $z \to W_{Si}$), defines a reduction of V_t of the top channel near the corner(s). (Note that if w_{Si} were large, the 1-D V_t of the top channel (with $E_z = 0$) would be proportional to qN_Bh_{Si} [Lim83]. Also note that our explanation of diminishing V_t versus z could be given analogously in terms of V_t versus x of the side channels.) The lower V_t in the corner regions is consistent with thinking of these regions as quasi-1-D symmetrical DGFETs having an effective body thickness ($t_{Si(eff)}$) less than w_{Si}, where V_t depends on $qN_Bt_{ox}t_{Si(eff)}/2\varepsilon_{ox}$ [Lim83]. Such thinking is also consistent with the finding in [Doy03b] that the corner devices show reduced DIBL since, in FDFET and DGFET, the SCEs tend to be suppressed as t_{Si} is scaled down [Tri03].

Because of technological and electrical problems associated with controlled heavy doping in ultra-thin-silicon bodies/channels, viable nonclassical devices will, more than likely, be undoped, and will rely on metal gates with specific work functions (Φ_G) for V_t control [Tri03]. Further, we surmise that low N_B in (5.1) could influence the noted corner effects. So, we use MEDICI to simulate the same SOI cross-section of Figure 5.3, but with $N_B = 0$ (For the MEDICI simulations, we set $N_B = 1.0x10^{10}cm^{-3}$ to avoid numerical instabilities. Such a low N_B ($< \sim10^{15}cm^{-3}$) is virtually equivalent to $N_B = 0$ with regard to the subthreshold characteristics [Tri04].), which, for a scaled device with a small-volume body/channel, will be the actual doping density in the undoped device. We assume a midgap metal gate ($\Phi_G = 4.6V$). Predicted n(x, z) versus V_G for this device is shown in Figure 5.4. The bulk V_t is a bit higher than that of Figure 5.3, but note that the low-V_G corner effects have been suppressed! For high V_G (strong inversion), there is some electron build-up in the corner regions, but, as in Figure 5.3, it is relatively insignificant, in accord with results in [Doy03b]. In this case, even though $n > N_B$, (5.1) with $N_B = 0$ is still accurate for weak inversion because the inversion-charge



(a)

(b)

Figure 5.4 MEDICI-predicted electron concentration of an undoped TG nMOSFET. (a) 2D electron distribution n(x, z'), where z' = z + 100nm, in the silicon-body/channel cross section A-A' (Figure 5.1(c)); $h_{Si} = w_{Si}$ = 30nm, V_G = 0.4V. (b) Predicted electron density n(x=0, z) across the top surface of the cross section for various values of V_G. Note, for this device with N_B = 0, the absence of any n build-up in the corner regions for low V_G (subthreshold region). The electron build-up for high V_G, like that in Figure 5.3, is not significant.

(neglected in (5.1)) contribution to V_t is negligible. Thus, as implied by the 1-D versions of (5.1), both the E_x and E_z partial derivatives are relatively small, and hence $N_{B(eff)}$ is virtually zero and there is no significant reduction in V_t as the corners are approached, either in z or in x. The nearly uniform n(x, z) in Figure 5.4(a), which reflects uniform electric potential in the silicon body, supports this explanation of no corner effects for $N_B = 0$. And, the subthreshold $n(V_G)$ variation in Figure 5.4(b) reflects ideal gate swing, S = 60mV/decade at T = 300K. Additional MEDICI simulations of TG structures with h_{Si} and/or w_{Si} increased to 60nm also showed no corner effects as long as N_B was low. We note that such elimination of the corner effects will prevail even for finite N_B defined by natural (unintentional) doping of larger bodies/channels.

In our 2-D analysis of the corner effects here, we have ignored SCEs (3-D effects), which tend to reduce V_t, but which tend to be suppressed by scaling the silicon body dimensions [Tri03]. So, for short L_g, the bulk and corner V_t's will be lower, but, thinking of the corner devices as symmetrical DGFET with relatively thin $t_{Si(eff)}$ as previously discussed, we can see that the reduction of the bulk V_t will exceed that of the corner V_t. Hence, for short L_g, the corner effects in the high-N_B device will be less discernible than in Figure 5.3. However, if the TGFET is properly scaled to control SCEs, this amelioration of the corner effects will be less significant. The need for ad hoc suppression of them remains, and hence the general use of undoped bodies to do so is proposed. This proposal, and our design insights regarding scaled TGFET, are supported by the 3-D simulations of corner effects in scaled FinFETs and shallow-trench-isolation (STI) bulk-silicon MOSFETs reported in [Bur03].

5.3 Short-Channel Effects of Scaled TGFETs

With the insights of previous section, we examine the feasibility of the TGFET for nanoscale CMOS applications with regard to SCEs and layout area. We check the SCEs for varying dimensions of the body, with and without doping, using Davinci [Dav03].

Following [Doy03a], we first assume that a reasonable TGFET design could, with reference to Figure 5.1(c), be $h_{Si} = w_{Si} = L_{eff}$ (assumed to equal L_g). For L_{eff} = 28nm devices, we let the gate-oxide thickness (t_{ox} = EOT) be 1.1nm and the buried-oxide thickness (t_{BOX}) be 200nm. For the 3-D device simulations, we assume abrupt source/drain junctions with 10nm gate overlaps (L_g = 48nm with L_{eff} = 28nm being the metallurgical channel length), meaning that fringing fields outside the intrinsic device are negligible. Figure 5.5 shows Davinci-predicted I_{DS}-V_{GS} characteristics of the TG nMOSFET, with both doped and undoped bodies. For the doped device, $N_B = 8.0\text{x}10^{18}\text{cm}^{-3}$ and the gate material is n^+ polysilicon. This device exhibits excellent subthreshold characteristics; the drain-induced barrier lowering (DIBL) is 35mV/V and the subthreshold slope, or swing factor (S) is 75mV, both of which are comparable to corresponding results in [Doy03a]. However, as we have previously explained in previous section, the good SCE control of the doped TGFET results from the

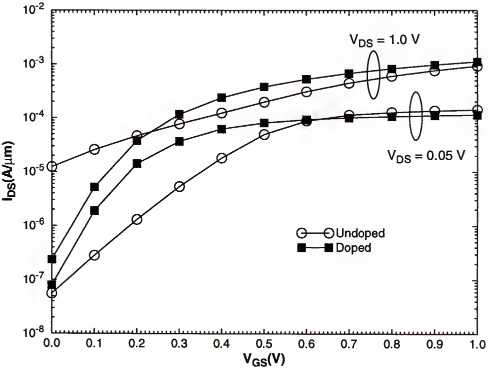


Figure 5.5 Davinci-predicted current-voltage characteristics of doped (N_B = 8.0x10^{18}cm^{-3} with n^+ polysilicon gate) and undoped (midgap gate) n-channel TGFETs; t_{ox} = 1.1nm, t_{BOX} = 200nm, and $h_{Si} = w_{Si} = L_{eff}$ = 28nm.

predominant subthreshold current flowing in the corner regions of the body, which can be thought of as DG devices with very thin effective body thickness, $t_{Si(eff)} < w_{Si}$.

To examine more closely the doped-body TGFET, we performed 3-D numerical simulations of devices with varying body dimensions. As shown in Figure 5.6, the predicted subthreshold characteristics are virtually independent of the body dimensions, which solidifies our insight that the subthreshold characteristics are governed by the corner regions having lower V_t than that of bulk-body device. Note in Figure 5.6, however, that the strong-inversion channel current is proportional to the effective gate width ($2h_{Si} + w_{Si}$) of each device, reflecting three surface channels with higher conductance than the corners.

The noted insights concerning the body-doped TGFET seem to imply an optimal device. Corner conduction in weak inversion yields good control of SCEs and relatively low off-state current (I_{off}), whereas the three surface channels in strong inversion imply good on-state current (I_{on}), as shown in [Doy03a] and [Doy03b], albeit without the insights noted here. In an actual TGFET, however, $t_{Si(eff)}$ depends on the finite radius of curvature of the corners [Doy03b], which could be difficult to control. Further, controlled doping of nanoscale silicon bodies is virtually impossible; the fluctuation in the number of channel dopants and their spacial randomness render unacceptable variations in V_t [Xio03]. Hence, the doped TGFET design is not technologically feasible.

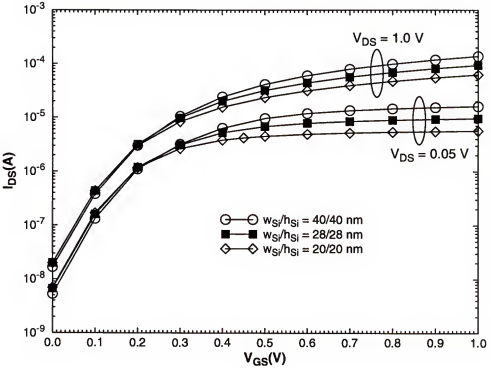


Figure 5.6 Davinci-predicted current-voltage characteristics of doped-body ($N_B = 8.0\text{x}10^{18}\text{cm}^{-3}$ with n^+ polysilicon gate) n-channel TGFETs; t_{ox} = 1.1nm, t_{BOX} = 200nm, and L_{eff} = 28nm. (The numerical simulation of the largest-body device was done using larger mesh-point spacings, which could yield slightly erroneous solutions in the corner regions and hence underlie the noticeably lower current in the subthreshold region.)

For the undoped ($N_B < \sim 1x10^{16}cm^{-3}$) TGFET in Figure 5.5, we assumed a midgap metal gate ($\Phi_G = 4.6V$) for threshold control. Note that the predicted subthreshold characteristics are much worse than those of the doped device (and are prohibitive), even though both devices have the same body dimensions. In this device, the corner conduction is suppressed as discussed in previous section, and the subthreshold current flows throughout the entire body. Obviously, however, the body is too large to control the SCEs. Nonetheless, the undoped body is the viable design, and hence the body dimensions needed to suppress the SCEs must be checked.

To check SCE control in the undoped-body TGFET, we performed 3-D simulations of the basic structure in Figure 5.1, still assuming $\Phi_G = 4.6V$, $t_{ox} = 1.1nm$, $t_{BOX} = 200nm$, and abrupt source/drain-body junctions. Several sets of body dimensions were tried in the simulations to obtain reasonable subthreshold characteristics, i.e., DIBL = 100mV/V and S = 80mV. Required combinations of h_{Si} and w_{Si} for the TGFET with $L_{eff} = 28nm$ are shown in Figure 5.7; in fact, the dimensions needed are about the same for both the DIBL and S specifications. Due to absence of the corner effects now, the body dimensions needed to suppress the SCEs are much smaller than those intimated in [Doy03a], e.g., $h_{Si} = w_{Si} = L_{eff}$, indicated by the 'x' in Fig. 4, which yields the poor subthreshold characteristics shown in Figure 5.5. Note also the extreme cases: for large h_{Si}, the results give the required thin w_{Si} for the DG FinFET ($w_{Si(DG)} = 0.53L_{eff}$) to get the specified DIBL and S; and for large w_{Si}, they give the

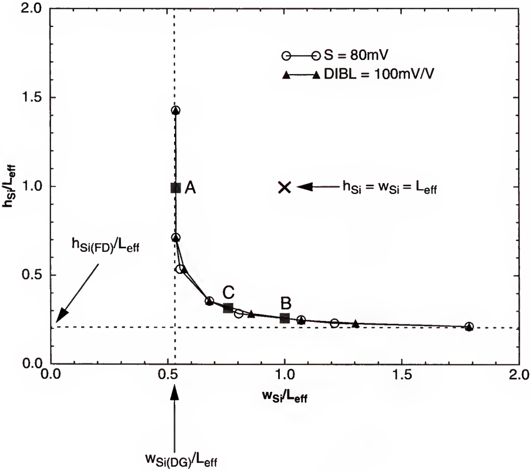


Figure 5.7 Davinci-predicted TGFET body dimensions needed (below and left of the curves) to suppress SCEs as noted; L_{eff} = 28nm, t_{ox} = 1.1nm, V_{DD} = 1.0V. For the extreme cases, $w_{Si(DG)}$ and $h_{Si(FD)}$ represent the required UTB thinness of the DG FinFET and the planar FDFET, respectively, to suppress the SCEs. The points A and B represent the body dimensions for $h_{Si} = L_{eff}$ and $w_{Si} = L_{eff}$, respectively, showing that TGFETs have negligible scaling advantage over the DG FinFET and the FDFET. Point C is an "optimal" TGFET design, having w_{Si} and h_{Si} equal to about 1.4 times $w_{Si(DG)}$ and $h_{Si(FD)}$, respectively. Note that the design with $h_{Si} = w_{Si} = L_{eff}$ (point ×) is clearly not feasible.

required thin h_{Si} for the FDFET ($h_{Si(FD)} = 0.21L_{eff}$). These UTB requirements are consistent with results in [Tri03] and [Fos03b].

To stress the much more stringent body scaling needed for the nanoscale TGFETs relative to that ($h_{Si} \sim w_{Si} \sim L_{eff}$) implied in [Doy03a] and [Doy03b], we note in Figure 5.7 (point A) that with $h_{Si} = L_{eff}$, $w_{Si} = w_{Si(DG)}$ is required to suppress the SCEs. This means that the TGFET has no scaling advantage over the DG FinFET at all! Clearly then, h_{Si} should be made as high as possible to maximize the effective gate width, creating in essence the DG FinFET with gate width $\sim 2h_{Si}$. Also, we note in Figure 5.7 (point B) that with $w_{Si} = L_{eff}$, the required h_{Si} for the TGFET is less than 20% larger than that needed for the FDFET ($h_{Si(FD)}$). This means that the TGFET has very little scaling advantage over the FDFET!

Our results and conclusions regarding the TGFET versus the FDFET are inconsistent with recently reported experimental results [Kri03], which suggest that TGFETs with $w_{Si}/L_g > 2.8$ show excellent subthreshold characteristics down to $L_g \sim 25$nm, even for $h_{Si}/L_g \sim 0.4$. Our results in Figure 5.7 show that $h_{Si}/L_g < 0.2$ would be needed to get the results reported in [Kri03]. Whereas there could be uncertainty in L_{eff} vs. L_g, and/or in t_{Si} ($= h_{Si}$), in [Kri03], the authors suggest that local strain in the body due to the metal (NiSi) gate and/or the SOI mesa isolation benefits the SCE control (as well as carrier mobility and I_{on}). Strained TGFETs might then be feasible, but understanding and controlling the strain constitute a formidable task.

5.4 Layout Area

With the required body dimensions for SCE control given in the preceding section, we now examine the CMOS layout-area issue of TGFETs, and compare it with those of planar FDFETs and DG FinFETs. We focus on gate areas, assuming that the peripheral areas of the three devices scale proportionally. Multi-finger structures as illustrated in Figure 5.8, with their pertinent dimensions, will generally be required. The 2-D cross-sectional view A-A' of the multi-gate/finger transistor, indicated in Figure 5.8 (a), is shown in Figure 5.8 (b) to relate the effective gate width (W_{eff}) of each finger (per pitch in the layout) to the pitch (P). In order to calculate the gate layout areas of the TGFET and the DG FinFET needed to achieve the same total drive current, we assume I_{on} per unit width of each device is the same, and equal to that of the FDFET, which will be used as a reference. With this assumption, gate layout areas for multi-gate/finger TGFETs and DG FinFETs are expressed as

$$A_{TGFET} = L_g\left[\frac{W_g}{2h_{Si} + w_{Si}}P\right] \tag{5.2}$$

$$A_{FinFET} = L_g\left[\frac{W_g}{2h_{Si}}P\right] \tag{5.3}$$

where W_g is the total gate width needed (which defines the number of fingers). Note that (5.2) and (5.3) depend on W_{eff} for each device; i.e., as indicated in Figure 5.8 (b), $W_{eff} = 2h_{Si} + w_{Si}$ for the TGFET and $W_{eff} = 2h_{Si}$ for the DG FinFET. For the planar FDFET, the gate layout area (A_{FD}) is simply W_gL_g, and $W_{eff} = P$.

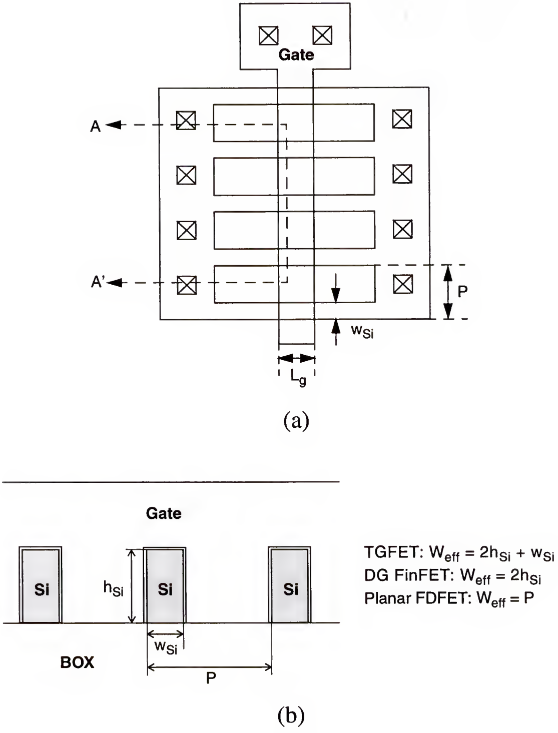

Gate
A
A'
P
w_{Si}
L_g
(a)
Gate
Si
h_{Si}
Si
Si
w_{Si}
BOX
P
TGFET: $W_{eff} = 2h_{Si} + w_{Si}$
DG FinFET: $W_{eff} = 2h_{Si}$
Planar FDFET: $W_{eff} = P$
(b)

Figure 5.8 The schematic illustrations of the multi-gate/finger transistor. (a) Top view. (b) Cross-sectional view A-A' as indicated in (a). The effective gate width of each finger (per pitch in the layout) is noted for the TGFET, the DG FinFET, and the planar FDFET in (b).

Based on the simulation results in Figure 5.7, we assume, for SCE control, $w_{Si}/L_{eff} = 0.5$ for the DG FinFET and $h_{Si}/L_{eff} = 0.2$ for the FDFET. The purported advantage of the TGFET is the relaxation of the silicon-body dimensions required, relative to the DG FinFET and the FDFET. However, this advantage is severely restricted due to SCEs as illustrated in Figure 5.7. In order to relax w_{Si} and h_{Si} at the same rate while maintaining the same SCEs, the body dimensions of an "optimal" TGFET are chosen to be about 1.4 times $w_{Si(DG)}$ and $h_{Si(FD)}$, as indicated by point C in Figure 5.7. Then, using (5.2) and (5.3), we show in Figure 5.9 calculated ratios (R_a) of A_{TGFET} and A_{FinFET} to A_{FD} versus L_g (assumed to equal L_{eff}) at different technology nodes, with all body dimensions set as noted for SCE control. For the calculations, L_g and P were obtained from the ITRS [Sem01] CMOS projections for the high-performance (HP) logic technology and the low standby-power (LSTP) technology. We assumed maximum aspect ratios ($R_f = h_{Si}/w_{Si}$) of 5 for the DG FinFET. Since R_a of the DG FinFET is inversely proportional to R_f, it would increase in Figure 5.9 by a factor of 5/3 ~ 5/4 for smaller $R_f = 3 \sim 4$, which is perhaps more technologically reasonable for the scaled device.

As evident in Figure 5.9, R_a of the TGFET is much higher than that of the DG FinFET, which can be <1 for higher R_f. At the $L_g = 37$nm node of the HP logic technology, the TGFET would require almost four times the gate area required for the DG FinFET. This severe layout-area disadvantage prevails, for the same body-dimension factor of 1.4, all the

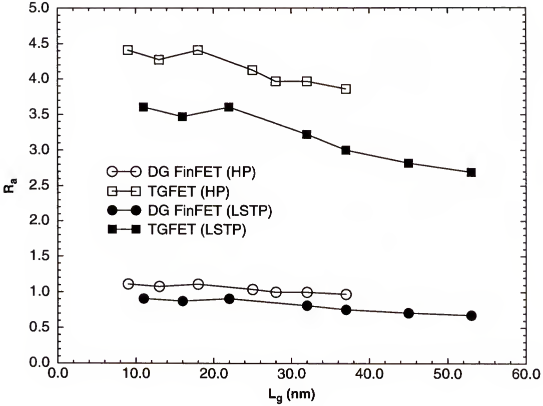


Figure 5.9 Calculated layout-area ratios (R_a) of the TGFET with $w_{Si} = 1.4w_{Si(DG)}$ and $h_{Si} = 1.4h_{Si(FD)}$ and the DG FinFET with $R_f = h_{Si}/w_{Si} = 5$ relative to the planar FDFET, versus gate length; L_g (= L_{eff}) and P for the calculations were obtained from the ITRS projections for HP and LSTP CMOS technologies [Sem01].

way to the end of the ITRS where $L_g \sim 10$nm. Note also in Figure 5.9 that R_a is better for both devices in the LSTP technology because, at a given node for specified P, L_g is scaled less aggressively. (For a specific L_g, the HP vs. LSTP differences in Figure 5.9 reflect different values of P.) However, the severe TGFET layout-area disadvantage remains. Finally, we note that all the ratios plotted in Figure 5.9 would be reduced by a factor of two if the multi-gate/finger devices were defined by spacer lithography [Doy03b], thereby substantially enhancing the relative layout-area advantage of DG FinFETs.

The layout-area comparisons in Figure 5.9 reflect the smaller W_{eff} for the TGFET structure used. To check its optimization, the 3-D simulation results for S = 80mV in the TGFET with L_{eff} = 28nm in Figure 5.7 are replotted in Figure 5.10. Lines A and B there represent $W_{eff} = P$ for the TGFET and DG FinFET, respectively, as for the planar FDFET. To get W_{eff} equal to or wider than P, the respective body dimensions must be above the lines A and B. For the TGFET, our "optimal" design (point C in Figure 5.7) is well below line A. We could redesign it as indicated in Figure 5.10 for wider W_{eff} and smaller R_a. However, it is evident that moving toward line A with the same SCE control would require either decreasing w_{Si} (path ① in Figure 5.10), which yields a virtual DG FinFET (W_{eff} of the TGFET is wider than that of the DG FinFET having the same body dimensions due to active top gate, but the added gate width is relatively small when $h_{Si}/w_{Si} >> 1$.), or decreasing h_{Si} (path ② in Figure 5.10), which

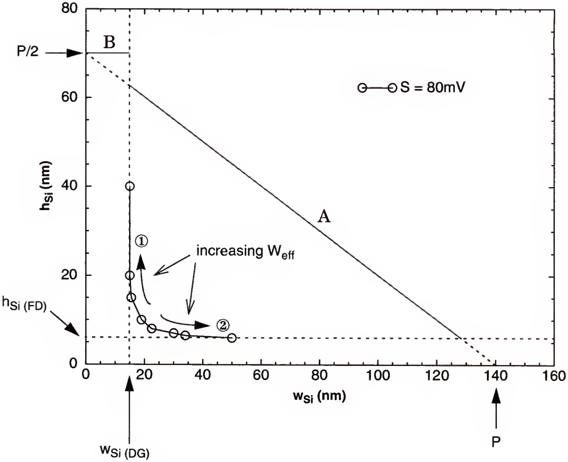


Figure 5.10 Davinci-predicted body dimensions of the TGFET needed for $S = 80mV$; $L_{eff} = 28nm$, $t_{ox} = 1.1nm$, and $V_{DD} = 1.0V$, with lines A and B representing $W_{eff} = P$ for the TGFET and the DG FinFET, respectively. The arrows indicate possible design-optimization paths, but lead to DG FinFET and planar FDFET structures.

yields a virtual FDFET. The TGFET structure is inferior, unless it is transformed into a DG FinFET or a FDFET! Note also in Figure 5.10 that the DG FinFET would be made more area-efficient than the FDFET, i.e., have a wider W_{eff}, if h_{Si} could be above line B; this would require $R_f > 5$.

From the 3-D device simulation results in Figure 5.7, we have noted the requirement of body thickness for SCE control to be $w_{Si}/L_{eff} \sim 0.5$ for the DG FinFET and $h_{Si}/L_{eff} \sim 0.2$ for the FDFET. Hence, if we set the pragmatic lower limit of body thickness (w_{Si} or h_{Si}) as 5nm due to manufacturing burden and quantization effects from the structural confinement [Tri03], we note that the FDFET cannot be scaled down below $L_g \sim$ 25nm (where we are still assuming $L_{eff} \sim L_g$.). However, the DG FinFET can be scaled down to $L_g \sim$ 10nm. Indeed, this superior scalability of the undoped-body DG FinFET seems to render it the most promising candidate for future nanoscale CMOS.

5.5 Conclusions

Via 2-D numerical device simulations and quasi-2-D analyses of nanoscale TGFETs, we have shown and explained that the reduced V_t of the corner regions in the body/channel can be eliminated by leaving the body undoped, and hence relying on a metal gate with proper work function for V_t control. This finding adds to the technological and electrical reasons [Tri03] for proposing the use of undoped bodies in nonclassical MOSFETs; the problem of controlling the shape of the corners in the TGFET [Doy03b], which is probably not possible, is eliminated.

Focusing on viable undoped bodies, we examined the UTB dimensions needed to control SCEs in nanoscale TGFETs by 3-D numerical simulations. The results showed that much more stringent body scaling is needed for undoped TGFETs, which have no corner effects, relative to the doped ones, which are technologically and electrically infeasible. When the undoped-body dimensions are scaled for adequate SCE control, we find that the TGFET suffers from a significant layout-area disadvantage relative to the DG FinFET and the planar FDFET. We hence conclude that the TGFET structure, without some benefit from strain [Kri03], is not feasible because, when adequately scaled for SCE control, its gate layout-area efficiency is relatively poor and can be improved only by evolving it into a DG FinFET with $h_{Si}/w_{Si} >> 1$ or a FDFET with $h_{Si}/w_{Si} << 1$. Further, our insights imply that this layout disadvantage is not peculiar only to the TGFET but is characteristic of any multi-gate MOSFET with more than two gates, for example, the gate-all-around device [Col90] and the surrounding-gate transistor [Tak91]. We thus argue that the DG FinFET, which is more scalable than the FDFET, is the nonclassical device with the most potential for future nanoscale CMOS applications.

CHAPTER 6
SUMMARY AND SUGGESTIONS FOR FUTURE WORK

6.1 Summary

In this dissertation, analysis and modeling of parasitic effects in advanced SOI CMOS technologies, including UTB transistors and TGFETs, were presented, and the feasibility of the TGFET was discussed. The major contributions of the research are summarized as follows.

In chapter 2, gate-to-body tunneling current (I_{GB}) in contemporary MOSFETs was physically modeled and implemented in the process-based compact model UFPDB. The valence-band electron tunneling (J_{VBE}) to the n^+-polysilicon gate of the nMOSFET, predominant for inversion conditions, and the conduction-band electron tunneling (J_{CBE}) to the substrate, predominant for depletion/accumulation conditions, were modeled by modifying the classical independent-electron formalism. For J_{VBE}, several scaling effects (i.e., quantization effect in the channel, exchange energy of inversion electrons, and Fermi-Dirac statistics in the high-doped gate), which have been overlooked in contemporary modeling of gate direct-tunneling current, were identified and accounted for. The composite I_{GB} model, applicable to pMOSFET as well, showed good agreement with measured data for gate-oxide thicknesses varying down to 1.65nm.

In chapter 3, using the I_{GB} model in UFPDB/Spice3, we showed how the tunneling current tends to suppress dynamic floating-body effects in some PD/SOI circuits, providing benefits that follow from the complex interactions among the bias conditions, circuit topologies, and switching patterns. However, the significance of these benefits of I_{GB}, as well as its detrimental effects [Fun00, Jos01, Chu02, Poi02], tend to be diminished in dynamic operation of scaled PD/SOI CMOS circuits because the supply voltage will be reduced and the scaled SOI technology will necessarily be optimized to suppress the detrimental DC floating-body effects on off-state current.

In chapter 4, the effects of nonclassical undoped UTB device structure on GIDL current were examined. GIDL current is enhanced by the needed higher gate work function of the UTB device and amplified by parasitic BJT. To predict the amplification of the GIDL current as well as transient BJT current, the parasitic BJT model for nonclassical UTB transistor was upgraded and implemented in the UFDG MOSFET models. The model was verified with the results of MEDICI simulations and experimental data of FDFETs and symmetrical DGFETs, showing good agreement for various device dimensions and reasonable bias conditions. For a dynamic logic circuit, we showed that the parasitic BJT can cause an erroneous evaluation for some bias conditions and switching patterns due to inherent low gate capacitance. Thus, a rigorous simulation is needed to ensure right operation and reduce the noise margin for nonclassical UTB transistor circuits via a reliable parasitic BJT model dependent on the UTB device structure.

In chapter 5, we showed and explained that the reduced V_t of the corner regions in the body/channel of triple-gate FinFETs (TGFETs) can be eliminated by leaving the body undoped, and hence relying on a metal gate with proper work function for V_t control. This finding adds to the technological and electrical reasons [Tri03] for proposing the use of undoped bodies in nonclassical MOSFETs; the problem of controlling the shape of the corners in the TGFET [Doy03b], which is probably not possible, was eliminated. Focusing on viable undoped bodies, we examined the UTB dimensions needed to control SCEs in nanoscale TGFETs by 3-D numerical simulations. The results showed that much more stringent body scaling is needed for undoped TGFETs, which have no corner effects, relative to the doped ones, which are technologically and electrically infeasible. When the undoped-body dimensions are scaled for adequate SCE control, we found that the TGFET suffers from a significant layout-area disadvantage relative to the DG FinFET and the planar FDFET. We hence conclude that the TGFET structure, without some benefit from strain [Kri03], is not feasible because, when adequately scaled for SCE control, its gate layout-area efficiency is relatively poor and can be improved only by evolving it into a DG FinFET with $h_{Si}/w_{Si} >> 1$ or a FDFET with $h_{Si}/w_{Si} << 1$. Further, our insights imply that this layout disadvantage is not peculiar only to the TGFET but is characteristic of any multi-gate MOSFET with more than two gates, for example, the gate-all-around device [Col90] and the surrounding-gate transistor [Tak91]. We thus argue that the DG FinFET, which is more scalable than the FDFET, is the nonclassical device with the most potential for future nanoscale CMOS applications.

6.2 Suggestions for Future Work

The following research tasks are suggested as future work regarding advanced SOI devices.

For a high-K dielectric, the gate-to-channel tunneling current would still be considered as a major leakage current. Modeling and analysis of the gate-to-channel tunneling current are suggested for UTB transistors, with (or without) high-K dielectrics.

Recently, a study based on simulation shows that a gate-underlap structure would be unavoidable and needed for nonclassical UTB transistors. However, even for the gate-underlap structure, a numerical simulation-based study shows GIDL current, which can be induced by the gate fringing field, can still limit the off-currents for low-standby power devices. Therefore, the modeling of GIDL current in nonclassical UTB transistors employing the gate-underlap structure is recommended.

2-D device simulation shows that unamplified GIDL current depends on UTB thickness in symmetrical DGFETs, i.e., GIDL current increases for thinner bodies. This dependence perhaps results from volume inversion in the thin Si film under the gate-drain overlap. Modeling of GIDL current in UTB devices is thus recommended.

APPENDIX
DERIVATION OF THE INTERBAND TUNNELING CURRENT EQUATION

Previous studies of direct tunneling currents in MOS devices have focused primarily on the component which flows between the gate and the channel, i.e., intraband (conduction band-to-conduction band) tunneling. However, modeling of the tunneling current that flows between the gate and the body (as in chapter 2), i.e., interband (valence band-to-conduction band) tunneling, has not been adequately documented. In the interband tunneling, the proper characterization of the tunneling current is somewhat different from the commonly used direct tunneling current expression [Sha99, Dep95, She72] since the electron density-of-states effective mass is different in the valence- and conduction-bands and the Si bandgap is indirect, which implies the involvement of phonons in the tunneling process [Mol64]. However, we do not explicitly account for the conservation of transverse momentum (via phonons) in this case; in fact, the functional dependence of tunneling probability on electric field is almost identical [Mol64] to that for a direct bandgap semiconductor, even though the needed phonon involvement decreases tunneling probability dramatically [Mol64]. With this insight, we develop here a general expression for both interband and intraband tunneling current by generalizing the classical model of Harrison [Har61].

The general probability per unit time of the tunneling transition of an electron in a state **a** on one side of the tunneling region to a state **b** on the other side was given as [Har61]

$$P_{ab} = \frac{2\pi}{\hbar}|M_{ab}|^2\rho_b f_a(1-f_b) \tag{A.1}$$

$|M_{ab}|$: the matrix element for the transition

ρ_b: the density of states at **b**

f_a, f_b: the probabilities of occupation of the states **a** and **b**, respectively.

The same formalism applies to the reverse tunneling process:

$$P_{ba} = \frac{2\pi}{\hbar}|M_{ba}|^2\rho_a f_b(1-f_a). \tag{A.2}$$

The total tunneling current is derived by summing over all states **a** of fixed k_t (transverse momentum), summing over k_t, multiplying by 2 for spin, and multiplying by the electronic charge q:

$$j_{ab} = 2q\sum_{k_{ta}}\int_0^\infty P_{ab}\rho_a dE_x \tag{A.3}$$

$$j_{ba} = 2q\sum_{k_{tb}}\int_0^\infty P_{ba}\rho_b dE_x. \tag{A.4}$$

The net current is then

$$j = j_{ab} - j_{ba} = 2q\left(\sum_{k_{ta}}\int_0^\infty P_{ab}\rho_a dE_x - \sum_{k_{tb}}\int_0^\infty P_{ba}\rho_b dE_x\right). \tag{A.5}$$

By replacing the sum over k_t by an integral over a constant energy surface, i.e.,

$$\sum_{k_t} \Rightarrow \int \frac{1}{\left(\frac{2\pi}{L}\right)^2} d^2k_t = L^2 \int \frac{1}{(2\pi)^2} 2\pi k_t dk_t = L^2 \int \frac{m_t^*}{2\pi\hbar^2} dE_t \tag{A.6}$$

which follows from

$$E_t = \frac{\hbar^2 k_t^2}{2m_t^*}$$

for 2D electrons, we have, per unit area (L^2),

$$j_{a\rightarrow b} = 2q \frac{1}{2\pi\hbar^2}\left[m_{ta}^* \int_0^{E_{t(max)}} \int_0^{\infty} P_{ab}\rho_a dE_x dE_t - m_{tb}^* \int_0^{E_{t(max)}} \int_0^{\infty} P_{ba}\rho_b dE_x dE_t\right]$$

$$= \frac{2q}{\hbar^3}\left[m_{ta}^* \int_0^{E_{t(max)}} \int_0^{\infty} |M_{ab}|^2 \rho_b \rho_a f_a (1 - f_b) dE_x dE_t\right.$$

$$\left. - m_{tb}^* \int_0^{E_{t(max)}} \int_0^{\infty} |M_{ba}|^2 \rho_a \rho_b f_b (1 - f_a) dE_x dE_t\right]. \tag{A.7}$$

With Harrison's formalism [Har61] for the matrix element in (A.7),

$$|M_{ab}|^2 \rho_a \rho_b = \frac{1}{4\pi^2} e^{-\eta} \tag{A.8}$$

where $\eta = 2\int_0^{tox} |k_x| dx$ (with k_x being the momentum perpendicular to the SiO_2-Si interface),

(A.7) becomes

$$j_{a\rightarrow b} = \frac{4\pi q}{h^3}\left[m_{ta}^* \int_0^{E_{t(max)}} \int_0^{\infty} e^{-\eta} f_a (1 - f_b) dE_x dE_t\right.$$

$$\left. - m_{tb}^* \int_0^{E_{t(max)}} \int_0^{\infty} e^{-\eta} f_b (1 - f_a) dE_x dE_t\right] \tag{A.9}$$

which is a general (interband or intraband) tunneling equation.

REFERENCE LIST

[Bur03] A. Burenkov and J. Lorenz, "On the Role of Corner Effect in FinFETs," Proc. European Workshop on Ultimate Integration of Silicon, Udine, Italy, pp. 31-34, Mar. 2003.

[Cha02] R. Chau, B. Doyle, J. Kavalieros, D. Barlage, A. Murthy, M. Doczy, R. Rios, T. Linton, R. Arghavani, B. Jin, S. Datta, and S. Hareland, "Advanced Depleted-Substrate Transistors: Single-Gate, Double-Gate, and Tri-Gate," Proc. Internat. Conf. on Solid State Devices and Materials," Nagoya, Japan, pp. 68-69, Sept. 2002.

[Che88] C.-E. D. Chen, M. Matloubian, R. Sundaresan, B.-Y. Mao, C. C. Wei, and G. P. Pollack, "Single-Transistor Latch in SOI MOSFET's," IEEE Electron Devices Lett., vol. 9, pp. 636-638, Dec. 1988.

[Che92] J. Chen, F. Assaderaghi, P. K. Ko, and C. Hu, "The Enhancement of Gate-Induced-Drain-Leakage (GIDL) Current in Short-Channel SOI MOSFET and Its Application in Measuring Lateral Bipolar Current Gain β," IEEE Electron Devices Lett., vol. 13, pp. 572-574, Nov. 1992.

[Chi96] S. C. Chin, Y.-C. Tseng, and J. C. S. Woo, "Parasitic Bipolar Turn-On of PD-SOI MOSFETs in Dynamic Logic Circuits," Proc. IEEE Internat. SOI Conf., Sanibel Island, FL, pp. 144-145, Oct. 1996.

[Chi01] M.-H. Chiang, Process-Based Compact Modeling and Analysis of Silicon-on-Insulator CMOS Devices and Circuits, Inclusing Double-Gate MOSFETs, Ph.D. Dissertation. Gainesville: Univ. of Florida, 2001.

[Cho91] J.-Y. Choi and J. G. Fossum, "Analysis and Control of Floating-Body Bipolar Effects in Fully Depleted Submicrometer SOI MOSFET's," IEEE Trans. Electron Devices, vol. 38, pp. 1384-1391, Jun. 1991.

[Chu97] C. T. Chuang, P. F. Lu, J. Ji, L. F. Wagner, S. Chu, and C. J. Anderson, "Dual-Mode Parasitic Bipolar Effect in Dynamic CVSL XOR Circuit with Floating-Body Partially-Depleted SOI Devices," Proc. Symp. VLSI Tech., Syst., and Application, Taipei, Taiwan, pp. 288-292, 1997.

[Chu98] C. T. Chuang, "Design Considerations of SOI Digital CMOS VLSI," Proc. IEEE Internat. SOI Conf., Stuart, FL, pp. 5-8, Oct. 1998.

[Chu02] C. T. Chuang and R. Puri, "Effects of Gate-to-Body Tunneling Current on Pass-Transistor Based PD/SOI CMOS Circuits," Proc. IEEE Internat. SOI Conf., Williamsburg, VA, pp. 121-122, Oct. 2002.

[Col90] J. P. Colinge, M. H. Gao, A. Romano-Rodríguez, H. Maes, and C. Claeys, "Silicon-on-Insulator 'Gate-All-Around Device,' " IEDM Tech. Dig., pp. 595-598, Dec. 1990.

[Dav03] Davinci-2003.06 User Guide. Durham, NC: Synopsys, Inc., 2003.

[Dep95] M. Depas, B. Vermeire, P. W. Mertens, R. L. Van Meirhaeghe, and M. M. Heyns, "Determination of Tunnelling Parameters in Ultra-Thin Oxide Layer Poly-Si/SiO2/Si Structure," Solid-State Electron., vol. 38, pp. 1465-1471, Aug. 1995.

[Dor02] B. Doris, M. Ieong, T. Kanarsky, Y. Zhang, R. A. Roy, O. Dokumaci, Z. Ren, F.-F. Jamin, L. Shi, W. Natzle, H.-J. Huang, J. Mezzapelle, A. Mocuta, S. Womack, M. Gribelyuk, E. C. Jones, R. J. Miller, H.-S. P. Wong, and W. Haensch, "Extreme Scaling with Ultra-Thin Si Channel MOSFETs," IEDM Tech. Dig., pp. 267-270, Dec. 2002.

[Doy03a] B. Doyle, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalierous, T. Linton, A. Murthy, R. Rios, and R. Chau, "High Performance Fully-Depleted Tri-Gate CMOS Transistors," IEEE Electron Devices Lett., vol. 24, pp. 263-265, Apr. 2003.

[Doy03b] B. Doyle, B. Boyanov, S. Datta, M. Doczy, S. Hareland, B. Jin, J. Kavalierous, T. Linton, R. Rios, and R. Chau, "Tri-Gate Fully-Depleted CMOS Transistors: Fabrication, Design and Layout," Proc. VLSI Tech. Symp., Kyoto, Japan, pp. 133-134, Jun. 2003.

[Fos98] J. G. Fossum, M.-H. Chiang, and T. W. Houston, "Design Issues and Insights for Low-Voltage High-Density SOI DRAM," IEEE Trans. Electron Devices, vol. 45, pp. 1055-1062, May 1998

[Fos99] J. G. Fossum, K. Kim, and Y. Chong, "Extremely Scaled Double-Gate CMOS Performance Projections, Including GIDL-Controlled Off-State Current," IEEE Trans. Electron Devices, vol. 46, pp. 2195-2200, Nov. 1999

[Fos02a] J. G. Fossum, UFSOI MOSFET Models (Ver. 7.5) Including UFPDB-2.5 User's Guide. Gainesville, FL: Univ. of Florida, 2002.

[Fos02b] J. G. Fossum, L. Ge, and M.-H. Chiang, "Speed Superiority of Scaled Double-Gate CMOS," IEEE Trans. Electron Devices, vol. 49, pp. 808-811, May 2002.

[Fos03a] J. G. Fossum, UFDG MOSFET Model (Ver. 2.3) User's Guide. Gainesville, FL: Univ. of Florida, 2003.

[Fos03b] J. G. Fossum, M. M. Chowdhury, V. P. Trivedi, T.-J. King, Y.-K. Choi, J. An, and B. Yu, "Physical Insights on Design and Modeling of Nanoscale FinFETs," IEDM Tech. Dig., pp. 679-682, Dec. 2003.

[Fos04] J. G. Fossum, L. Ge, M.-H. Chiang, V. P. Trivedi, M. M. Chowdhury, L. Mathew, G. O. Workman, and B.-Y. Nguyen, "A Process/Physics-Based Compact Model for Nonclassical CMOS Device and Circuit Design," Solid-State Electron., vol. 48, pp. 919-926, Jun. 2004.

[Fra01] D. J. Frank, R. H. Dennard, E. Nowak, P. M. Solomon, Y. Taur, and H.-S. P. Wong, "Device Scaling Limits of Si MOSFETs and Their Application Dependencies," Proc. IEEE, vol. 89, pp. 259-288, Mar. 2001

[Fun00] S. K. H. Fung, N. Zamdmer, P. J. OLdiges, J. Sleight, A. Mocuta, M. Sherony, S.-H. Lo, R. Joshi, C. T. Chuang, I. Yang, S. Crowder, T. C. Chen, F. Assaderaghi, and G. Shahidi, "Controlling Floating-Body Effects for 0.13mm and 0.10mm SOI CMOS," IEDM Tech. Dig., pp. 231-234, Dec. 2000.

[Ge02] L. Ge and J. G. Fossum, "Analytical Modeling of Quantization and Volume Inversion in Thin Si-Film Double-Gate MOSFETs," IEEE Trans. Electron Devices, vol. 49, pp. 287-294, Feb. 2002.

[Har61] W. A. Harrison, "Tunneling from an Independent-Particle Point of View," Phys. Rev., vol. 123, pp. 85-89, July 1961.

[Jos01] R. V. Joshi, C. T. Chuang, S. K. H. Fung, F. Assaderaghi, M. Sherony, I. Yang, and G. Shahidi, "Effects of Gate-to-Body Tunneling Current on PD/SOI CMOS SRAM," Proc. VLSI Tech. Symp., Kyoto, Japan, pp. 75-76, June 2001.

[Ked01] J. Kedzierski, D. M. Fried, E. J. Nowak, T. Kanarsky, J. H. Rankin, H. Hanafi, W. Natzle, D. Boyd, Y. Zhang, R. A. Roy, J. Newbury, C. Yu, Q. Yang, P. Saunders, C. P. Willets, A. Johnson, S. P. Cole, H. E. Young, N. Carpenter, D. Rakowski, B. A. Rainey, P. E. Cottrell, M. Ieong, and H.-S. P. Wong, "High-Performance Symmetric-Gate and CMOS-Compatible Vt Asymmetric-Gate FinFET Devices," IEDM Tech. Dig., pp. 437-440, Dec. 2001.

[Ked02] J. Kedzierski, E. J. Nowak, T. Kanarsky, Y. Zhang, D. Boyd, R. Carruthers, C. Cabral, R. Amos, C. Lavoie, R. Roy, J. Newbury, E. Sullivan, J. Benedict, P. Saunders, K. Wong, D. Canaperi, M. Krishnan, K.-L. Lee, B. A. Rainey, D. Fried, P. Cottrell, H.-S. P. Wong, M. Ieong, and W. Haensch, "Metal-Gate FinFET and Fully-Depleted SOI Devices Using Total Gate Silicidation," IEDM Tech. Dig., pp. 247-250, Dec. 2002.

[Kha00] Khairurrijal, W. Mizubayashi, S. Miyazaki, and M. Hirose, "Analytic Model of Direct Tunnel Current through Ultrathin Gate Oxides," J. Appl. Phys., vol. 87, pp. 3000-3005, Mar. 2000.

[Kim01] Keunwoo Kim, Design and Analysis of Double-Gate CMOS for Low-Voltage Integrated Circuit Applications, Including Physical Modeling of Silicon-on-Insulator MOSFETs, Ph.D. Dissertation. Gainesville: Univ. of Florida, 2001.

[Kri96] Srinath Krishnan, Analysis and Modeling of Nonlocal and Dynamic Floating-Body Effects for Application in Scaled SOI CMOS Technology, Ph.D. Dissertation. Gainesville: Univ. of Florida, 1996.

[Kri02] Z. Krivokapic, W. Maszara, K. Achutan, P. King, J. Gray, M. Sidorow, E. Zhao, J. Zhang, J. Chan, A. Marathe, and M.-R. Lin, "Nickel Silicide Metal Gate FDSOI Devices with Improved Gate Oxide Leakage," IEDM Tech. Dig., pp. 271-274, Dec. 2002.

[Kri03] Z. Krivokapic, V. Moroz, W. Maszara, and M.-R. Lin, "Locally Strained Ultra-Thin Channel 25nm Narrow FDSOI Devices with Metal Gate and Mesa Isolation," IEDM Tech. Dig., pp. 445-448, Dec. 2003.

[Lee00] W.-C. Lee and C. Hu, "Modeling Gate and Substrate Currents due to Conduction- and Valence-Band Electron and Hole Tunneling," Proc. VLSI Tech. Symp., Honolulu, HI, pp. 198-199, June 2000.

[Lim83] H.-K. Lim and J. G. Fossum, "Threshold Voltage of Thin-Film Silicon-on-Insulator (SOI) MOSFETs," IEEE Trans. Electron Devices, vol. 49, pp. 1244-1251, Oct. 1983.

[Lo97] S. H. Lo, D. A. Buchanan, Y. Taur, and W. Wang, "Quantum-Mechanical Modeling of Electron Tunneling Current from the Inversion Layer of Ultra-Thin-Oxide nMOSFET's," IEEE Electron Devices Lett., vol. 18, pp. 209-211, May 1997.

[Lu97] P.-F. Lu, C.-T. Chuang, J. Ji, L. F. Wagner, C.-M. Hsieh, J. B. Kuang, L. L.-C. Hsu, M. M. Pelella, S.-F. S. Chu, and C. J. Anderson, "Floating-Body Effects in Partially Depleted SOI CMOS Circuits," IEEE J. Solid-State Circuits, vol. 32, pp. 1241-1253, Aug. 1997.

[Maj00] B. Majkusiak and M. H. Badri, "Semiconductor Thickness and Back-Gate Voltage Effects on the Gate Tunnel Current in the MOS/SOI System with an Ultrathin Oxide," IEEE Trans. Electron Devices, vol. 47, pp. 2347-2351, Dec. 2000.

[Med99] MEDICI-4.0 User's Manual. Fremont, CA: Avant! Corp., 1999.

[Mol64] J. L. Moll, Physics of Semiconductors. New York: McGraw-Hill Book Co., 1964.

[Mul85] R. S. Muller, and T. I. Kamins, Device Electronics for Integrated Circuits. New York: Wiley, 1985.

[Nil78] N. G. Nilsson, "Empirical Approximations for the Fermi Energy in a Semiconductor with Parabolic Bands," Appl. Phys. Lett., vol. 33, pp. 653-654, Oct. 1978.

[Pel95] M. M. Pelella, J. G. Fossum, D. Suh, S. Krishnan, and K. A. Jenkins, "Low-Voltage Transient Bipolar Effect Induced by Dynamic Floating-Body Charging in PD/SOI MOSFETs," Proc. IEEE Internat. SOI Conf., Tucson, AZ, pp. 8-9, Oct. 1995.

[Pel02] M. M. Pelella and J. G. Fossum, "On the Performance Advantage of PD/SOI CMOS With Floating Bodies," IEEE Trans. Electron Devices, vol. 49, pp. 96-104, Jan. 2002.

[Poi02] T. Poiroux, O. Faynot, C. Tabone, H. Tigelaar, H. Mogul, N. Bresson, and S. Cristoloveanu, "Emerging Floating-Body Effects in Advanced Partially-Depleted SOI Devices," Proc. IEEE Internat. SOI Conf., Williamsburg, VA, pp. 99-100, Oct. 2002.

[Reg99] L. F. Register, E. Rosenbaum, and K. Yang, "Analytic Model for Direct Tunneling Current in Polycrystalline Silicon-Gate Metal-Oxide-Semiconductor Devices," Appl. Phys. Lett., vol. 74, pp. 457-459, Jan. 1999.

[Sah91] C. T. Sah, Fundamentals of Solid-State Electronics. Singapore: World Scientific Publishing Co., 1991.

[Sak03] W. Sakamoto, T. Endoh, H. Sakuraba, and F. Masuoka, "Reduction of Pass-Gate Leakage by Silicon-Thickness Thinning in Double-Gate MOSFETs," Electrochemical Society Proc., Paris, France, vol. 2003-05, pp. 331-336, Apr. 2003.

[Sch04] T. Schulz, C. Pacha, R. J. Luyken, M. Stadele, J. Hartwich, L. Dreeskornfeld, E. Landgraf, J. Kretz, W. Rosner, M. Specht, F. Hofmann, and L. Risch, "Impact of Technology Parameters on Device Performance of UTB-SOI CMOS," Solid-State Electron., vol. 48, pp. 521-527, Apr. 2004.

[Sem01] International Technology Roadmap for Semiconductors. San Jose, CA: Semiconductor Industry Assoc., 2001.

[Sha99] A. Shanware, J. P. Shiely H. Z. Massoud, E. Vogel, K. Henson, A. Srivastava, C. Osburn, J. R. Hauser, and J. J. Wortman, "Extraction of the Gate Oxide Thickness of N-and P-Channel MOSFETs Below 20Å from the Substrate Current Resulting from Valence-Band Electron Tunneling," IEDM Tech. Dig., pp. 815-818, Dec. 1999.

[She72] J. Shewchun and V. A. K. Temple, "Theoretical Tunneling Current Characteristics of the SIS Diode," J. Appl. Phys., vol. 43, pp. 5051-5061, Dec. 1972.

[Shi98] Y. Shi, T. P. Ma, S. Prasad, and S. Dhanda, "Polarity Dependent Gate Tunneling Currents in Dual-Gate CMOSFET's," IEEE Trans. Electron Devices, vol. 45, pp. 2355-2360, Nov. 1998.

[Shi99] Y. Shi, X. Wang, and T.-P. Ma, "Electrical Properties of High-Quality Ultrathin Nitrde/Oxide Stack Dielectrics," IEEE Trans. Electron Devices, vol. 46, pp. 362-368, Feb. 1999

[Sim84] J. G. Simmons, F. L. Hsueh, and L. Faraone, "Two-Carrier Conduction in MOS Tunnel-Oxides II-Theory," Solid-State Electron., vol. 27, pp. 1131-1139, 1984.

[Sin02] S. Sinha, M.-H. Chiang, and M. M. Pelella, "Impact of Gate Tunneling on the Nature of the Charge Dump Current in 100nm PDSOI CMOS," Proc. IEEE Internat. SOI Conf., Williamsburg, VA, pp. 41-42, Oct. 2002.

[Slo91] J. W. Slotboom, G. Streutker, M. J. v. Dort, P. H. Woerlee, A. Pruijmboom, and D. J. Gravesteijn, "Non-Local Impact Ionization in Silicon Devices," IEDM Tech. Dig., pp. 127-130, Dec. 1991.

[Ste73] F. Stern, "Electron Exchange Energy in Si Inversion Layers," Phys. Rev. Lett., vol. 30, No. 7, pp. 278-280, Feb. 1973.

[Sze81] S. M. Sze, Physics of semiconductor devices, 2nd ed. New York: Wiley, 1981.

[Tak91] H. Takato, K. Sunouchi, N. Okabe, A. Nitayama, K. Hieda, F. Horiguchi, and F. Masuoka, "Impact of Surrounding Gate Transistor (SGT) for Ultra-High-Density LSI's," IEEE Trans. Electron Devices, vol. 38, pp. 573-578, Mar. 1991.

[Tau98a] Y. Taur and T. H. Ning, Fundamentals of Modern VLSI Devices. New York: Cambridge University Press, 1998.

[Tri03] V. P. Trivedi and J. G. Fossum, "Scaling Fully Depleted SOI CMOS," IEEE Trans. Electron Devices, vol. 50, pp. 2095-2103, Oct. 2003.

[Tri04] V. P. Trivedi and J. G. Fossum, "Non-Classical CMOS Devices with Undoped Bodies: Electrostatics and the Depletion Approximation," unpublished, 2004.

[Van94] M. J. Van Dort, P. H. Woerlee and A. J. Walker, "A Simple Model for Quantization Effects in Heavily-Doped Silicon MOSFETs at Inversion Conditions," Solid-State Electron., vol. 37, pp. 411-414, Mar. 1994.

[Wes93] N. Weste and K. Eshraghian, Principles of CMOS VLSI Design-A Systems Perspective, 2nd ed. Reading, MA: Addison-Wesley, 1993.

[Xio03] S. Xiong and J. Bokor, "Sensitivity of Double-Gate and FinFET Devices to Process Variations," IEEE Trans. Electron Devices, vol. 50, pp. 2255-2260, Nov. 2003.

[Yan99] N. Yang, W. K. Henson, J. R. Hauser, and J. J. Wortman, "Modeling Study of Ultrathin Gate Oxides Using Direct Tunneling Current and Capacitance-Voltage Measurements in MOS Devices," IEEE Trans. Electron Devices, vol. 46, pp. 1464-1470, July 1999.

[Yan00] K.-N. Yang, H.-T. Huang, M.-C. Chang, C.-M. Chu, Y.-S. Chen, M.-J. Chen, Y.-M. Lin, M.-C. Yu, S. M. Jang, D. C. H. Yu, and M. S. Liang, "A Physical Model for Hole Direct Tunneling Current in P+ Poly-Gate PMOSFETs with Ultrathin Gate Oxides," IEEE Trans. Electron Devices, vol. 47, pp. 2161-2166, Nov. 2000.

[Yeh95] P. C. Yeh and J. G. Fossum, "Physical Subthreshold MOSFET Modeling Applied to Viable Design of Deep-Submicrometer Fully Depleted SOI Low-Voltage CMOS Technology," IEEE Trans. Electron Devices, vol. 42, pp. 1605-1613, Sep. 1995.

[You88] K. K. Young and J. A. Burns, "Avalanch-Induced Drain-Source Breakdown in Silicon-on-Insulator n-MOSFET's," IEEE Trans. Electron Devices, vol. 35, pp. 426-431, Apr. 1988.

BIOGRAPHICAL SKETCH

Ji-Woon Yang was born in Kwangju, Korea. He received the B.S. and M.S. degrees in electrical engineering from the Korea University, Seoul, Korea, in 1991 and 1995, respectively.

From 1995 to 1999, he was a member of the research staff of the advanced device development group at the Hyundai Electronics Industry Co., Ichon, Korea, working on process integration, qualification, design, and optimization of deep sub-micron MOSFETs and SOI devices for 1Gb DRAM. He and his team developed the world-first fully functional 1Gb SOI DRAM in 1997. Since 1999, he has been pursuing a Ph.D. degree as a graduate research assistant at the University of Florida. His research interests involve device design, analysis, and modeling of high-scaled CMOS technology including fully depleted SOI transistor and double-gate FinFET.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Jerry G. Fossum, Chair
Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Gijs Bosman
Professor of Electrical and
Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William R. Eisenstadt

William R. Eisenstadt
Associate Professor of Electrical
and Computer Engineering

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Timothy J. Anderson

Timothy J. Anderson
Professor of Chemical Engineering

This dissertation was submitted to the Graduate Faculty of the College of Engineering and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August 2004

Pramod Khargonekar

Pramod P. Khargonekar
Dean, College of Engineering

Kenneth J. Gerhardt
Interim Dean, Graduate School